



Universidade do Estado do Rio de Janeiro

Centro de Ciências Sociais

Faculdade de Administração e Finanças

Tiago Rodrigues Guimarães

**Análise dos fatores que contribuem para a recuperação dos créditos de
ICMS inscritos em Dívida ativa no Estado do Rio de Janeiro**

Rio de Janeiro

2023

Tiago Rodrigues Guimarães

Análise dos fatores que contribuem para a recuperação dos créditos de ICMS inscritos em Dívida ativa no Estado do Rio de Janeiro

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Controladoria e Gestão Pública, da Universidade do Estado do Rio de Janeiro.

Orientador: Prof. Dr. José Francisco Moreira Pessanha

Rio de Janeiro

2023

CATALOGAÇÃO NA FONTE
UERJ/REDE SIRIUS/BIBLIOTECA CCS/B

G963 Guimaraes, Tiago Rodrigues
Análise dos fatores que contribuem para a recuperação dos créditos de ICMS inscritos em Dívida ativa no Estado do Rio de Janeiro / Tiago Rodrigues Guimaraes – 2023.
75 f.

Orientador: José Francisco Moreira Pessanha.
Dissertação (mestrado) – Universidade do Estado do Rio de Janeiro, Faculdade de Administração e Finanças.
Bibliografia: f.68-71.

1. Dívida ativa – Rio de Janeiro (Estado) - Teses. 2. Administração pública – Teses. I. Pessanha, José Francisco Moreira. II. Universidade do Estado do Rio de Janeiro. Faculdade de Administração e Finanças. III. Título.

CDU 336.2 (815.3)

Bibliotecária: Lucia Andrade – CRB7 / 5272

Autorizo, apenas para fins acadêmicos e científicos, a reprodução total ou parcial desta dissertação.

Assinatura

Data

Tiago Rodrigues Guimarães

**Análise dos fatores que contribuem para a recuperação dos créditos de ICMS inscritos
em Dívida ativa no Estado do Rio de Janeiro**

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Controladoria e Gestão Pública da Universidade do Estado do Rio de Janeiro.

Aprovado em 28 de julho de 2023:

Banca Examinadora:

Prof. Dr. José Francisco Moreira Pessanha (Orientador)
Faculdade de Administração e Finanças – UERJ

Prof. Dr. Francisco José dos Santos Alves
Faculdade de Administração e Finanças – UERJ

Prof. Dr. Reinaldo Castro Souza
Pontifícia Universidade Católica do Rio de Janeiro

Rio de Janeiro
2023

AGRADECIMENTOS

Ao meu orientador, José Francisco Moreira Pessanha, pela disponibilidade durante o processo de construção da dissertação e conclusão do curso de mestrado. Ao professor, Francisco José dos Santos Alves, pelo apoio técnico e suas valiosas contribuições.

À minha mãe, Lourdimar Rodrigues Santos, por todo o seu amor e suporte à minha educação.

Ao meu amigo Wallace da Silva Rocha pelo incentivo e estímulo a me inscrever no mestrado.

Aos Procuradores do Estado Hugo Wilken Maurell e Natália Faria de Souza pelo apoio e encorajamento para a conclusão do curso.

Aos meus amigos Ane Loize e Jonatas por toda ajuda, bem como, aos amigos da Procuradoria da Dívida Ativa que me incentivaram no processo de elaboração desse trabalho.

Às minhas amigas de turma, Glória e Úrsula, por sempre me ajudarem quando precisei.

Ao Programa de Pós-Graduação em Controladoria e Gestão Pública da UERJ, por toda a atenção ao corpo discente e a todos os servidores pelo comprometimento e esforços empreendidos para manter a possível normalidade dos estudos e pesquisas, especialmente no primeiro ano de pandemia da Covid-19.

A Deus, pela saúde e por sua graça.

A vida só pode ser compreendida olhando-se para trás, mas só pode ser vivida olhando-se para a frente.

Soren Kierkegaard.

RESUMO

GUIMARÃES, Tiago Rodrigues. **Análise dos fatores que contribuem para a recuperação dos créditos de ICMS inscritos em Dívida ativa no Estado do Rio de Janeiro**. 2023: 76f. Dissertação (Mestrado em Controladoria e Gestão Pública) – Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2023.

As Procuradorias, em geral, são conhecidas por atuar de maneira burocrática, seguindo formalmente os atos legalmente previstos, porém sem avaliar a eficiência e a efetividade das ações empreendidas. Isso resulta no acúmulo de execuções fiscais no judiciário, gerando um impacto negativo nos cofres públicos, pois muitas vezes não há um retorno proporcional aos esforços realizados. Diante desse contexto, o objetivo deste trabalho é contribuir para o aprimoramento da gestão da dívida ativa do Estado do Rio de Janeiro. Para isso, propõe-se uma abordagem que fornecerá uma classificação da probabilidade da recuperabilidade dos créditos inscritos, permitindo que a Procuradoria Geral do Estado do Rio de Janeiro (PGE/RJ) direcione seus esforços de recuperação de forma mais eficiente e, ao mesmo tempo, reconheça contabilmente esses créditos de maneira mais objetiva. A metodologia adotada neste estudo envolve a análise quantitativa e o uso do algoritmo de aprendizado de máquina conhecido como XGBoost. Através dessa abordagem, serão identificados os fatores que influenciam a probabilidade de recuperação das Certidões de Dívida Ativa (CDA), levando em consideração variáveis como quantidade de dias em débito, valor total da dívida, dias de empresa, situação cadastral e região, entre outras. Com base nos resultados obtidos, será estabelecido um *rating* de créditos, proporcionando uma visão mais clara sobre a probabilidade de recuperação de cada CDA e permitindo a priorização dos esforços de recuperação nos casos com maior probabilidade de sucesso. Além disso, esses resultados poderão contribuir para uma melhor estimativa de perdas relacionadas à dívida ativa do Estado do Rio de Janeiro. Atualmente, a PGE/RJ utiliza um método baseado no julgamento profissional dos procuradores, sem a utilização de critérios estatísticos objetivos, para estimar as perdas com créditos da dívida ativa. No entanto, com base nos resultados do modelo proposto, é estimado que as perdas cheguem a aproximadamente 98,64% do valor total da dívida ativa, revelando a necessidade de uma revisão dessa metodologia. Portanto, os resultados obtidos por meio do XGBoost podem contribuir significativamente para a construção de uma estimativa mais precisa das perdas com os créditos da dívida ativa do Estado do Rio de Janeiro. Vale destacar que o modelo desenvolvido apresentou uma alta *performance* na previsão da recuperabilidade dos créditos. Além disso, esse enfoque mais objetivo e baseado em dados pode auxiliar a PGE/RJ a direcionar seus esforços de recuperação de forma mais eficaz, resultando em um melhor desempenho na gestão da dívida ativa e no retorno financeiro para os cofres públicos.

Palavras-Chaves: Dívida Ativa. Gestão Pública. XGBoost. Algoritmo de aprendizado de máquina. Classificação de créditos.

ABSTRACT

GUIMARÃES, Tiago Rodrigues. **Analysis of the factors that contribute to the recovery of ICMS credits enrolled in the Active Debt of the State of Rio de Janeiro.** 2023: 76f. Dissertação (Mestrado em Controladoria e Gestão Pública) – Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2023.

Prosecutor's Offices are generally known for their bureaucratic approach, following formal procedures as legally prescribed, but often without evaluating the efficiency and effectiveness of their actions. This leads to a backlog of tax executions in the judicial system, causing a negative impact on public funds as efforts are not always proportionally rewarded. In light of this context, the objective of this study is to contribute to the improvement of the management of the State of Rio de Janeiro's active debt. To achieve this, an approach is proposed that provides a classification of the probability of recovering the registered credits, allowing the State Prosecutor's Office of Rio de Janeiro (PGE/RJ) to allocate recovery efforts more efficiently while recognizing these credits more objectively in accounting terms. The methodology adopted in this study involves quantitative analysis and the use of the machine learning algorithm known as XGBoost. Through this approach, factors influencing the probability of recovering Active Debt Certificates (CDA) will be identified, considering variables such as days in arrears, total debt amount, days of company existence, registration status, and region, among others. Based on the obtained results, a credit rating will be established, providing a clearer view of the probability of recovery for each CDA and allowing for the prioritization of recovery efforts in cases with a higher likelihood of success. Additionally, these results can contribute to a better estimation of losses related to the State of Rio de Janeiro's active debt. Currently, PGE/RJ relies on a method based on professional judgment of prosecutors, without the use of objective statistical criteria, to estimate losses from active debt credits. However, based on the results of the proposed model, it is estimated that the losses amount to approximately 98,64% of the total active debt value, highlighting the need for a revision of this methodology. Therefore, the results obtained through the XGBoost can significantly contribute to a more accurate estimation of losses from the State of Rio de Janeiro's active debt credits. It is noteworthy that the developed model demonstrated high performance in predicting credit recoverability. Moreover, this more objective and data-driven approach can assist PGE/RJ in directing recovery efforts more effectively, resulting in improved management of the active debt and financial returns to public funds.

Keywords: Active Debt. Public Management. XGBoost. Machine Learning Algorithm. Credit Rating.

LISTA DE FIGURAS

Figura 1: Organograma da estrutura Jurídica da Dívida Ativa da PGE/RJ	29
Figura 2: Organograma da estrutura administrativa da Dívida Ativa da PGE/RJ.....	29
Figura 3: Modelo de curva ROC	43
Figura 4: Recuperação e liquidação da dívida ativa 2018 a 2022	48
Figura 5: Total do Estoque e percentual recuperado 2018 a 2022	49
Figura 6: Matriz de confusão.....	51
Figura 7: Curva ROC.....	51
Figura 8: Importância das variáveis.....	52

LISTA DE TABELAS

Tabela 1: Variáveis explicativas do estudo	37
Tabela 2: Matriz de confusão para classificador com duas classes	41
Tabela 3: <i>Rating</i> do potencial de recuperabilidade	44
Tabela 4: Composição do estoque da dívida ativa total ao longo dos anos (2018-2022).....	45
Tabela 5: Crescimento da dívida ativa ao longo dos anos (2018-2022).....	46
Tabela 6: Participação do ICMS no estoque total da dívida ativa ao longo dos anos (2018-2022).....	46
Tabela 7: Composição dos valores recuperados e liquidados da dívida ativa total ao longo dos anos (2018-2022).....	47
Tabela 8: Participação na recuperação e liquidação da dívida ativa por categoria ao longo dos anos (2018-2022).....	48
Tabela 9: Total do estoque da dívida ativa e percentual recuperado.....	49
Tabela 10: Faixa de Classificação (<i>Rating</i>).....	53
Tabela 11: Classificação das CDA por faixa de <i>rating</i> e média de dias de recuperabilidade ..	54
Tabela 12: Classificação das CDA por <i>rating</i> , quantidade de CDA, valor total e valor médio	55
Tabela 13: Classificação das CDA por <i>rating</i> e tempo médio de existência das empresas em dias.....	55
Tabela 14: Classificação das CDA por região em valores absolutos	56
Tabela 15: Classificação das CDA por débitos ajuizados	56
Tabela 16: Variáveis das Certidões de Dívida Ativa (CDA) - <i>rating</i> A (85,00% a 100,00%)	57
Tabela 17: Variáveis das Certidões de Dívida Ativa (CDA) referente ao <i>rating</i> B (70,00% a 84,99%)	58
Tabela 18: Variáveis das Certidões de Dívida Ativa (CDA) referente ao <i>rating</i> C (55,00% a 69,99%)	59
Tabela 19: Variáveis das Certidões de Dívida Ativa (CDA) referente ao <i>rating</i> D (40,00% a 54,99%)	60
Tabela 20: Variáveis das Certidões de Dívida Ativa (CDA) referente ao <i>rating</i> E (25,00% a 39,99%).....	61
Tabela 21: Variáveis das Certidões de Dívida Ativa (CDA) referente ao <i>rating</i> F (10,00% a 24,99%)	62
Tabela 22: Variáveis das Certidões de Dívida Ativa (CDA) referente ao <i>rating</i> G (0,01% a 9,99%)	63
Tabela 23: Estimativa de Perdas para CDA por Faixa de Classificação	65

LISTA DE QUADROS

Quadro 1: Estrutura Jurídica da Dívida Ativa da PGE-RJ	30
Quadro 2: Estrutura administrativa da Dívida Ativa da PGE-RJ	31

SUMÁRIO

	INTRODUÇÃO	13
1	REFERENCIAL TEÓRICO	18
1.1	Modelos de gestão aplicáveis ao contexto da administração tributária: perspectivas e desafios	18
1.2	Ciclo da dívida ativa	21
1.3	Contabilização da dívida ativa	22
1.4	Aprendizagem de máquina (<i>machine learning</i>)	25
1.5	Estrutura da procuradoria geral do estado do Rio de Janeiro	28
1.6	Análise da metodologia atual adotada para ajuste de perdas para dívida ativa no estado do Rio de Janeiro	32
2	METODOLOGIA	34
2.1	Identificação das bases de dados	34
2.2	Variáveis da pesquisa	35
2.3	Coleta e preparação dos dados	37
2.4	Análise dos dados	38
2.5	Rating do potencial de recuperabilidade da dívida ativa	44
3	DESCRIÇÃO E ANÁLISE DOS RESULTADOS	45
3.1	Análise descritiva da base de dados	45
3.2	Resultados do classificador <i>xgboost</i>	50
3.2.1	Faixa de classificação (<i>rating</i>) A.....	57
3.2.2	Faixa de classificação (<i>rating</i>) B	58
3.2.3	Faixa de classificação (<i>rating</i>) C	59
3.2.4	Faixa de classificação (<i>rating</i>) D.....	60
3.2.5	Faixa de classificação (<i>rating</i>) E	61
3.2.6	Faixa de classificação (<i>rating</i>) F	62
3.2.7	Faixa de classificação (<i>rating</i>) G.....	63
3.3	Proposta de modelo de ajuste de perdas para Dívida Ativa RJ	64
	CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS	67
	REFERÊNCIAS	69
	ANEXO – APURAÇÃO DA PERDA PELA PGE/RJ	73

INTRODUÇÃO

A dívida ativa é uma espécie de crédito público, cuja matéria é definida desde a Lei nº 4.320/64, sendo sua gestão econômica, orçamentária e financeira resultante de uma conjugação de critérios estabelecidos em diversos outros textos legais, abrange os créditos a favor da Fazenda Pública, cuja certeza e liquidez foram apuradas e que não tenham sido efetivamente recebidos nas datas aprazadas. É, portanto, uma fonte potencial de fluxos de caixa, com impacto financeiro positivo pela recuperação de valores espelhando créditos a receber.

Sob a ótica burocrática do Direito, a inscrição em dívida ativa e o ajuizamento de execuções fiscais são, tradicionalmente, entendidos como atos vinculados, podendo o agente público que deixar de praticá-los incorrer em responsabilidade funcional, cível e criminal.

Como consequência dessa visão tradicional, no Brasil, as Procuradorias tendem a atuar visando o cumprimento formal dos atos legalmente previstos, sem que haja avaliação numa perspectiva da eficiência e/ou efetividade das ações empreendidas, em um típico paradigma burocrático de gestão. O resultado tem sido o acúmulo de execuções fiscais junto ao judiciário, com prazo superior às demais classes processuais, sem que haja proporcional retorno aos cofres públicos.

Segundo o Relatório e Parecer Prévio do Tribunal de Contas do Estado do Rio de Janeiro (TCERJ) sobre as Contas do Governador do Estado do Rio de Janeiro referente ao exercício de 2021, o montante de créditos ainda não recuperados pelo Estado alcançou R\$121,98 bilhões no referido exercício, de maneira que o crescimento foi superior ao período anterior em R\$3,53 bilhões, sendo que, a maior contribuição para o estoque final da dívida ativa no exercício de 2021 advém da natureza ICMS, cujo valor total no respectivo ano foi de R\$ 118,09 bilhões (TCERJ, 2021).

Adicionalmente, o TCERJ (2021), por meio de Auditoria Financeira que contrapôs os valores apresentados em 2021 e o potencial de ingresso de recursos por intermédio da recuperação de créditos inadimplidos, constatou que, embora R\$40,48, bilhões dos créditos inscritos sejam considerados com remota probabilidade de recuperação, o referido montante não foi baixado contabilmente, revelando assim, uma fragilidade no controle contábil de tais valores.

Cabe registrar, ainda, que não obstante recente histórico de déficit financeiro apresentado pelo Governo do Estado Rio de Janeiro, cujos recursos arrecadados foram insuficientes frente às demandas orçamentárias, culminando em atrasos a fornecedores e servidores, o Estado apresentou no ano de 2022, o saldo total de mais de R\$135 bilhões de reais,

ficando atrás apenas do Estado de São Paulo no que diz respeito a créditos a receber em dívida ativa.

Ocorre que a efetiva arrecadação de todos os tributos está diretamente relacionada à responsabilidade na gestão fiscal, conforme preceitua o Art. 11 da Lei de Responsabilidade Fiscal (LRF) e que o crescimento do estoque da dívida ativa do Poder Público no decorrer dos anos, sem que haja controle robusto, tampouco a proporcional recuperação, demonstra a necessidade da adoção de medidas que promovam maior eficiência no monitoramento e cobrança dos créditos da dívida ativa.

Santos, Lima e Wilbert (2021) destacam que a dificuldade da recuperação desses valores pode ser o fato de até recentemente o reconhecimento da receita pública ser efetuado a partir de critérios orçamentários, ou seja, apenas no momento do efetivo ingresso do recurso, ficando para um segundo momento a contabilização dos valores pendentes de recebimento, ainda que cumprido seu fato gerador. Com o advento do processo de convergência da Contabilidade Pública brasileira aos padrões internacionais, o registro do direito a receber relativo a esses valores passou a ser efetuado por competência, com base no fato gerador, o que fez com que os recebíveis pudessem ser mais bem observados.

Nesse diapasão, convém destacar a classificação de crédito, ou *rating*, que é uma avaliação de risco feita por agências especializadas em relação à capacidade de um devedor honrar suas dívidas. Porém, a classificação de crédito da dívida ativa, que é a dívida que o governo tem a receber de pessoas físicas e jurídicas, não é tão usual quanto à classificação de crédito de empresas e governos que emitem títulos de dívida no mercado.

A avaliação da capacidade de pagamento da dívida ativa pode ser um fator determinante para verificar se é necessário fazer o reconhecimento das perdas desses créditos. Esse ajuste é importante para refletir a real situação dos créditos tributários, considerando a probabilidade de inadimplência e a capacidade financeira do devedor.

Uma empresa ou indivíduo que apresenta histórico de descumprimento de obrigações tributárias pode ter mais dificuldades em negociar o pagamento da dívida, o que pode resultar em um aumento do valor devido, em função da incidência de juros e multas. Por outro lado, um contribuinte que apresenta um bom histórico de cumprimento das obrigações tributárias pode ter mais facilidade em negociar o pagamento de débitos fiscais na dívida ativa.

Nesse contexto, a informação e o conhecimento combinados tornam-se ferramentas fundamentais para o estabelecimento, reposicionamento e/ou revisão de estratégias de gestão voltada para atuação ativa e eficiente, na medida em que possibilitam a transformação de dados *a priori* dispersos em base informacional estruturada aos tomadores de decisão (agentes

públicos), que poderão concentrar os esforços organizacionais nos casos com maiores probabilidades de recuperabilidade de créditos, assim como promover a adequada evidenciação de tais valores.

Considerando a constatada ineficiência do modelo tradicional de cobrança, aliada a necessidade de recuperação de créditos a fim de contribuir com incremento de recursos para orçamento público, assim como a necessidade de utilizar metodologia objetiva para calcular estimativa para o reconhecimento de ajustes dos valores para perdas na contabilidade, surge o seguinte questionamento: **quais fatores influenciam na recuperação dos créditos inscritos em dívida ativa no Estado do Rio de Janeiro?**

O presente trabalho objetiva aprofundar a análise e identificar, por meio de metodologia quantitativa, fatores, orientados pelos dados, que influenciem na recuperabilidade de créditos da dívida ativa, conforme expresso nas Certidões de Dívida Ativa (CDA). Para atingir esse objetivo, foram utilizados os dados de recuperação e estoque de dívida ativa dos créditos de ICMS inscritos no Estado do Rio de Janeiro no período de 2018 a 2022. Assim, tem-se como objetivos específicos:

1. Levantar o montante dos créditos recuperados e não recuperados nos últimos 5 anos;
2. Analisar a metodologia atual, adaptada pela PGE/RJ, para ajuste dos valores das perdas;
3. Elaborar modelo de classificação dos créditos (*rating*) que permita o direcionamento dos esforços da PGE/RJ na busca de créditos com maior probabilidade de recuperação;
4. Constituir uma classificação da probabilidade da recuperabilidade dos créditos inscritos em dívida ativa, permitindo o ajuste dos valores para perdas ou continuidade como provável recebimento.

Para realizar a classificação das CDA em faixas de classificação (*rating*), foi empregado o modelo XGBoost (CHEN e GUESTRIN, 2016). Esse modelo é uma técnica de aprendizado de máquina que permite treinar um algoritmo para aprender a classificar os dados com base em suas características. Neste estudo, foram utilizadas 10 variáveis para analisar o comportamento das CDA e identificar padrões relacionados à recuperação dos créditos. Essa abordagem visa fornecer uma compreensão mais completa e detalhada sobre as CDA, permitindo uma melhor estratégia de recuperação e evidenciação dos créditos.

Em face do atual cenário do Estado do Rio de Janeiro, o estudo sobre o modelo de classificação da dívida ativa se mostra ainda mais relevante e justificado. O Estado do Rio de Janeiro encontra-se em um regime de recuperação fiscal (RRF), desde 05/09/2017, data de homologação do acordo, que tem como objetivo reorganizar as finanças do Estado, equilibrar suas contas e permitir que o governo possa honrar seus compromissos financeiros.

No bojo do pedido apresentado ao Governo Federal há uma previsão de receita com créditos em dívida ativa, decorrentes de uma possível securitização dos créditos inscritos com base na Lei Estadual nº 7.040/2015.

Essa Lei ainda não foi aplicada de modo a atender a medida imposta pelo RRF, todavia, ao considerarmos o montante do estoque da dívida ativa, cujo valor total em 2022 superou R\$135 bilhões de reais, fica evidenciado a importância da recuperação dos créditos inscritos e não recuperados, pois, gera a expectativa de recebimentos futuros que promovam colaboração no equacionamento das dívidas do Estado.

Convém destacar que a gestão eficiente da dívida ativa representa uma fonte de recursos para o Estado, que pode ser utilizada para o pagamento de dívidas e para o equilíbrio das contas públicas. Além disso, o regime de recuperação fiscal exige que o Estado adote medidas de austeridade fiscal e redução de gastos, o que torna ainda mais crucial a eficiência na gestão da dívida ativa.

Diante disso, a realização do estudo sobre o modelo de classificação (*rating*) da dívida ativa no contexto do Estado do Rio de Janeiro se justifica pela necessidade de se buscar soluções mais eficientes e tecnológicas no contexto de um Estado que enfrenta sérios problemas fiscais e financeiros.

Igualmente, o estudo realizado pode desempenhar um papel importante na melhoria da fidedignidade das informações apresentadas no Balanço Geral do Estado. Ao utilizar modelos de *machine learning* para prever a probabilidade de retorno aos cofres públicos, com base em indicadores e algoritmos adequados, é possível obter estimativas mais precisas e condizentes com a realidade, permitindo uma melhor tomada de decisão e garantindo que as informações financeiras sejam mais confiáveis e precisas.

Ademais, o estudo poderá contribuir do ponto de vista prático, com governos e gestores no desenvolvimento de políticas institucionais de governança na área de gestão da dívida no que tange a melhoria do controle e cobrança de créditos da dívida ativa, bem como a sua evidenciação.

Outrossim, a baixa recuperabilidade dos créditos e a fragilidade da evidenciação contábil são indícios de que considerável parte das execuções fiscais são inefetivas devido à baixa eficácia do processo de cobrança de créditos em favor do Estado. Disso decorre a urgência do estudo de alternativas aos modelos atuais de recuperação desses créditos.

A pesquisa está estruturada da seguinte forma: esta introdução, que apresenta o problema de pesquisa, os objetivos e a justificativa do estudo. Em seguida, encontra-se a proposta de referencial teórico. Em seguida, estão detalhados os passos realizados na

metodologia e a análise dos resultados. Por fim, a são apresentadas as considerações finais e reflexões sobre a possibilidade de trabalhos futuros.

1 REFERENCIAL TEÓRICO

A Administração Pública é um tema de grande relevância para a sociedade, uma vez que esta é responsável por fornecer serviços públicos essenciais aos cidadãos. No Brasil, há dois modelos principais de administração pública: o modelo burocrático e o modelo gerencial. Enquanto o modelo burocrático é mais tradicional e valoriza a hierarquia e a rigidez de normas e procedimentos, o modelo gerencial busca uma gestão mais flexível, com ênfase em resultados e na participação dos cidadãos.

Um dos principais desafios enfrentados pela administração tributária é a cobrança da dívida ativa, que é o montante de tributos não pagos pelos contribuintes, mesmo após o prazo de vencimento. O ciclo da dívida ativa envolve desde a inscrição do débito na dívida ativa até a sua cobrança efetiva, passando por diversas etapas, como a notificação do contribuinte e a execução fiscal.

A contabilização da dívida ativa é um processo complexo, que envolve a estimativa do valor recuperável e o ajuste para perdas, de acordo com as normas contábeis aplicáveis. Além disso, a utilização de tecnologias como aprendizagem de máquina (ou *machine learning*) pode auxiliar no processo de tomada de decisão e na identificação de padrões, contribuindo para o aumento da eficiência e efetividade da administração tributária.

A seguir serão abordados esses tópicos contendo citação e resultado de pesquisas de autores seminais.

1.1 Modelos de gestão aplicáveis ao contexto da administração tributária: perspectivas e desafios

A administração tributária é responsável por gerenciar a arrecadação de impostos e tributos, bem como fiscalizar o cumprimento das obrigações fiscais pelos contribuintes. O desempenho eficiente dessas atividades é fundamental para garantir o funcionamento adequado do Estado e a oferta de serviços públicos de qualidade à população.

Nesse contexto, a administração pública pode ser dividida em dois modelos organizacionais reformistas: a burocrática e a gerencial (BRESSER-PEREIRA, 2022). A administração burocrática tem como principal característica a ênfase nas regras e procedimentos, enquanto a administração gerencial enfatiza a eficiência e a eficácia na prestação dos serviços públicos. Ambas as abordagens podem ser aplicadas na administração

tributária, mas cada uma delas apresenta vantagens e desvantagens em relação à gestão dos tributos.

É importante destacar que, embora tenham ocorrido iniciativas para reformar os modelos de gestão burocrática e gerencial da administração pública brasileira, o patrimonialismo ainda é uma realidade presente. Esse sistema coexiste com as tentativas de modernização e eficiência na gestão pública, e ainda representa um entrave para o desenvolvimento do país (BOAS et al., 2012).

O patrimonialismo é um sistema administrativo que se caracteriza pela ausência de distinção entre o público e o privado, em que as instituições públicas são utilizadas para atender aos interesses privados dos governantes. Segundo Boas et al. (2012), o patrimonialismo é um tipo de dominação tradicional, em que o governante exerce o poder como se fosse o dono do Estado e dos recursos públicos, e não há uma separação clara entre os interesses públicos e privados. Esse sistema é marcado pela corrupção, nepotismo e falta de profissionalismo na gestão pública.

A coexistência de tais modelos aponta para hibridização do modelo gerencial na dinâmica cultural brasileira e pode ser justificada pela transferência de tecnologia gerencial entre países centrais e periféricos, criando um efeito de “modernização superficial”, mas num contexto em que os problemas herdados pelo colonialismo ainda estão presentes nas organizações públicas do Brasil (ROSA e ALCADIPANI, 2013).

A substituição da gestão pública burocrática pela chamada administração pública gerencial com ênfase nos resultados teve como objetivo transformar organizações retirando do Estado o seu papel de executor primário e enfatizando o de formulador, fiscalizador e regulador de políticas públicas. A atualização do modelo visava ainda superar as “disfunções burocráticas” que dificultavam o desempenho político, econômico, administrativo e institucional do Estado brasileiro (LEITE, 2019).

A administração tributária burocrática é caracterizada pela estrutura hierárquica, controle de processos, especialização e regras claras e definidas. Esse modelo é orientado para a garantia do cumprimento das leis e normas tributárias, buscando a igualdade no tratamento dos contribuintes e a segurança jurídica. No entanto, esse modelo pode ser visto como engessado e limitante em relação à inovação e à eficiência na gestão tributária.

Por outro lado, a administração tributária gerencial tem como foco a eficiência na arrecadação e no controle dos tributos. Nesse modelo, há uma maior flexibilidade na gestão, com incentivo à inovação e à adoção de tecnologias para facilitar o cumprimento das obrigações tributárias. Além disso, a administração tributária gerencial busca estabelecer uma relação mais

próxima com os contribuintes, visando facilitar o cumprimento das obrigações tributárias e reduzir a sonegação.

Algumas pesquisas têm abordado o tema do desempenho na administração tributária, apresentando diferentes abordagens e metodologias de análise. Um exemplo é o estudo promovido pelo IPEA (2011), no qual verificou-se que apenas 25% da dívida ativa do governo federal é recuperada pela PGFN a um custo administrativo médio de R\$ 6.000,00 (dados de 2009) por processo e o tempo médio total de tramitação dos processos é de 9 anos, 9 meses e 16 dias.

Os resultados apontaram ainda que o *breaking even point*, isto é, o ponto a partir do qual é economicamente justificável promover-se judicialmente o executivo fiscal, é de R\$ 21.731,45. Ou seja, nas ações de execução fiscal de valor inferior a este, é improvável que a União consiga recuperar um valor igual ou superior ao custo do processamento judicial.

No entanto, a adoção da administração tributária gerencial pode ser vista como uma ameaça à segurança jurídica e à igualdade no tratamento dos contribuintes, uma vez que há uma maior discricionariedade na gestão tributária e um risco maior de arbitrariedade por parte das autoridades fiscais.

Por outro lado, ainda existem diversos desafios a serem enfrentados para eficiência da cobrança da dívida ativa, dentre os quais se destaca o lapso temporal no trâmite do processo administrativo fiscal que precede a inscrição em dívida ativa. Para Costa (2020) essa demora pode permitir que o devedor se organize para ocultar valores e evitar o pagamento da dívida, tornando-a potencialmente irrecuperável.

Oliveira e Santos (2013) propõem o uso da Lei de Improbidade Administrativa e eliminação das disfunções da burocracia como formas de aumentar a eficiência na cobrança da receita. Os autores também recomendam a criação de órgãos dedicados exclusivamente à receita pública e à cobrança da dívida ativa, para mudar a cultura da administração pública brasileira, que tende a focar mais no controle de despesas do que na gestão de receitas.

Portanto, a administração pública burocrática e gerencial no contexto da administração tributária deve ser entendida considerando as características peculiares da gestão tributária e os objetivos a serem alcançados, pois é importante que haja um equilíbrio entre os modelos para garantir a segurança jurídica.

Isto porque no contexto específico da cobrança da dívida ativa, a administração pública burocrática tende a ser mais rigorosa na cobrança dos tributos, seguindo rigidamente as normas e regras estabelecidas. Porém, esse modelo pode apresentar dificuldades na recuperação da

dívida ativa, principalmente quando há uma grande quantidade de processos a serem gerenciados.

Já a administração tributária gerencial tende a adotar uma abordagem mais flexível, utilizando estratégias como o parcelamento e a negociação da dívida, além de tecnologias que facilitem o acesso dos contribuintes às informações e ao cumprimento das obrigações tributárias. Isso pode trazer resultados positivos na recuperação da dívida ativa, mas também pode gerar críticas em relação à possibilidade de concessões indevidas e à falta de transparência na gestão tributária.

Dessa forma, é importante que a escolha entre a administração pública burocrática e gerencial no contexto da cobrança da dívida ativa seja realizada considerando a natureza específica dos tributos, o perfil dos contribuintes e os objetivos a serem alcançados. É necessário buscar um equilíbrio entre a rigidez necessária para garantir a recuperação dos tributos devidos e a flexibilidade suficiente para viabilizar a negociação e o diálogo com os contribuintes, visando uma maior eficiência na gestão tributária.

1.2 Ciclo da dívida ativa

O crédito tributário regularmente constituído pelo lançamento tem que ser quitado pelo sujeito passivo. Se, no entanto, não for efetuado o adimplemento voluntariamente pelo contribuinte, as Secretarias de Fazenda dos entes federativos remetem o crédito fiscal para as Procuradorias, para inscrição em dívida ativa, na forma do artigo 201 da Lei nº. 5.172/1966, o Código Tributário Nacional (CTN).

Uma importante compreensão relacionada à sistemática do ciclo da dívida ativa vem das etapas ou fases da receita pública. Rossi e Santos (2016) *apud* MCASP (2015) divide as receitas orçamentárias públicas em duas fases: planejamento e execução. Durante a fase de planejamento, ocorre a previsão das receitas, enquanto na fase de execução há o lançamento, a arrecadação e o recolhimento. Essas fases são cruciais para o controle das receitas e a gestão eficiente dos recursos públicos.

Os autores explicam que o lançamento é o procedimento administrativo que identifica o sujeito passivo e apura o montante do tributo devido. Já a arrecadação e o recolhimento são fases que se complementam e estão sujeitas ao regime de caixa. Quando há inadimplência do sujeito passivo, a Administração Pública realiza esforços para arrecadar e recolher a dívida. Após o prazo de pagamento, se a dívida não for quitada, ela é inscrita em registro próprio como dívida ativa.

Todavia, o CTN não estipula prazo contado a partir do vencimento, no qual o crédito terá que ser inscrito em dívida ativa. Dessa forma, compete à legislação tributária específica de cada Unidade da Federação determinar o intervalo de tempo entre o vencimento e a efetiva inscrição na dívida ativa.

No Estado do Rio de Janeiro, o § 6º do artigo 176 da Constituição Estadual e o inciso II do artigo 2º da Lei Complementar Estadual nº 15, de 25 de novembro de 1980 (com redação dada pela Lei Complementar nº 104/2002) determinam competência privativa da Procuradoria para promover a inscrição da dívida ativa do Estado e realizar a cobrança, quer de forma amigável quer de forma judicial.

A Procuradoria Geral do Estado do Rio de Janeiro (PGE-RJ) é o órgão responsável pela representação judicial e consultoria jurídica do Estado do Rio de Janeiro. Supervisiona os serviços jurídicos das administrações direta e indireta, atua no controle interno da legalidade dos atos da Administração Pública e defende judicial e extrajudicialmente os interesses legítimos do Estado.

A inscrição da dívida é realizada mediante “termo de inscrição da dívida ativa”, que instrumentaliza a Fazenda Pública com documento passível de execução. Não quer dizer, portanto que o poder público já esteja cobrando judicialmente o sujeito passivo.

Se ainda assim, o sujeito passivo não pagar, será emitida a Certidão de Inscrição em Dívida Ativa (CDA), título executivo extrajudicial, conforme artigo 784 do código de processo civil no inciso IX, pelo qual o Fisco passa o crédito para dívida ativa em fase de execução. O Art. 202 do CTN, juntamente com o artigo 195 do Decreto-Lei nº 05/1975, elencam os termos obrigatórios de inscrição e da Certidão da Dívida Ativa (CDA).

1.3 Contabilização da dívida ativa

De acordo com o Manual de Contabilidade Aplicado ao Setor Público - MCASP (2021), a dívida ativa é o crédito da Fazenda Pública, de natureza tributária ou não tributária, exigível pelo transcurso do prazo para pagamento, inscrito na forma da legislação própria após apurada a sua liquidez e certeza por órgão competente para tanto.

O Plano de Contas Aplicado ao Setor Público (PCASP), por seu turno, distingue a dívida ativa quanto à origem, conforme previsto na Lei nº 4.320/1964, em: **a)** dívida ativa tributária: é proveniente de obrigação legal relativa a tributos e respectivos adicionais e multas. e **b)** dívida ativa não tributária: é proveniente dos demais créditos da Fazenda Pública, decorrentes de contratos em geral ou de outras obrigações legais.

A inscrição do crédito em dívida ativa configura fato contábil permutativo, pois não altera o valor do patrimônio líquido do ente público. No órgão ou entidade de origem é baixado o crédito a receber contra uma Variação Patrimonial Diminutiva (VPD) e no órgão ou entidade competente para inscrição é reconhecido um crédito de dívida ativa contra uma Variação Patrimonial Aumentativa (VPA). Com relação à expectativa de realização, há troca do crédito a receber, registrado no curto prazo, pelo crédito de dívida ativa, evidenciado no longo prazo, tendo em vista que o inadimplemento torna incerto para realização do crédito.

Portanto, considerando-se o ente como um todo, a expectativa é que ocorra tão somente a troca do crédito a receber no curto prazo para o longo prazo (dívida ativa), sem alteração do valor do Patrimônio Líquido.

O MCASP (2021) dispõe, ainda, acerca de dois métodos de registro da dívida. O primeiro utiliza contas de controle para acompanhar todo o processo de inscrição da dívida ativa, desde o inadimplemento até a efetiva inscrição. Neste caso, é necessária maior integração entre as diversas etapas e unidades envolvidas. O segundo se restringe ao momento da efetiva inscrição dos valores em dívida ativa, dispensando o uso de contas de controle e de estrutura robusta de integração entre as áreas.

O Manual também alerta para necessidade de atualização monetária, juros, multas e outros encargos moratórios incidentes sobre os créditos inscritos em dívida ativa, os quais devem ser incorporados ao valor original inscrito, de acordo com o regime de competência.

Destaque-se que o ajuste para perdas de dívida ativa é um processo realizado pelas entidades públicas que possuem dívidas em aberto, a fim de contabilizar o valor que provavelmente não será recuperado. Esse processo é realizado em conformidade com as normas e princípios contábeis, a fim de garantir a transparência e a confiabilidade das informações financeiras.

O processo de ajuste para perdas de dívida ativa começa com a avaliação da situação da dívida, levando em consideração fatores como a idade da dívida, a situação econômica do devedor e as condições de pagamento acordadas. Com base nessa avaliação, é possível determinar uma estimativa do valor que provavelmente não será recuperado.

Em seguida, é feita a contabilização do ajuste para perdas, que é registrado como uma despesa no resultado do exercício. Esse registro deve seguir as normas e princípios contábeis, a fim de garantir que o valor registrado reflita adequadamente a realidade da situação.

É importante destacar que o ajuste para perdas de dívida ativa é um processo que deve ser realizado com cuidado e atenção, a fim de evitar distorções nas informações financeiras e garantir a transparência e a confiabilidade das informações contábeis. Por isso, é essencial que

as entidades públicas sigam as normas e princípios contábeis em vigor e contem com profissionais qualificados e experientes para realizar esse processo. Segundo Silva (2019) a política contábil de registro da Dívida ativa deve ser consistente, dando transparência a eventuais mudanças em metodologias que justifiquem a sua revisão.

Os créditos em dívida ativa possuem garantias jurídicas para sua cobrança, porém, frequentemente não são realizados devido a cancelamentos, prescrições e litígios judiciais. Por isso, é necessário registrar perdas esperadas por meio de uma conta redutora do ativo. A responsabilidade pelo cálculo e registro contábil do ajuste para perdas é do órgão ou entidade responsável pela gestão da dívida ativa.

O MCASP (2021) não estabelece uma metodologia específica para o cálculo do ajuste para perdas, considerando a diversidade da origem dos créditos e dos graus de estruturação das atividades de cobrança nos entes da Federação. Cada ente deve escolher a metodologia que melhor reflete a expectativa de recebimento dos créditos inscritos, e a metodologia utilizada e a memória de cálculo do ajuste para perdas devem ser divulgadas em Notas Explicativas.

A mensuração do ajuste para perdas deve ser baseada em estudos especializados que qualifiquem os créditos inscritos e evitem a superestimação ou subavaliação do patrimônio do ente público. Esses estudos podem considerar aspectos como o tipo de crédito (tributário ou não tributário), o prazo decorrido desde sua constituição e o andamento das ações de cobrança (extrajudicial ou judicial), entre outros.

É importante destacar que o MCASP (2021) estabelece que o ajuste para perdas de dívida ativa deve ser realizado anualmente, a fim de garantir a atualização das estimativas e a adequação das informações contábeis.

No que tange aos estudos acadêmicos, Rossi e Santos (2016) analisaram o processo de arrecadação e cobrança da dívida ativa na administração pública, incluindo os regimes contábeis aplicados para a receita pública, os métodos de gerenciamento do estoque de dívida ativa, as ações de controle e as políticas de renúncia de receita. Os autores se concentraram nas ações realizadas pelo Estado do Rio de Janeiro entre 2010 e 2014, usando abordagem quantitativa e métodos documental e bibliográfico. O estudo mostrou uma relação moderada entre as variáveis de provisão e registro de dívida ativa tributária e não tributária.

Outro estudo que merece destaque foi empreendido por Silva, Lima e Ferreira (2013), que analisou o tratamento contábil da dívida ativa tributária no Distrito Federal, no período de 2009 a 2012, à luz da teoria contábil. Foi constatado que há divergências em relação aos critérios de reconhecimento, mensuração e evidenciação da dívida ativa, com destaque para o critério de reconhecimento dos direitos a receber. Foi proposto que o registro do direito a

receber deve considerar o momento da caracterização do ativo como crédito tributário, e não a partir do fato gerador jurídico. Constatou-se ainda a necessidade de constituição de provisão para ajuste a valor recuperável da dívida ativa em dois momentos distintos. Sugere-se a realização de estudos sobre os critérios de mensuração para a constituição da provisão.

1.4 **Aprendizagem de máquina (*machine learning*)**

A aprendizagem de máquina (ou *machine learning*) é uma área da inteligência artificial que busca desenvolver algoritmos capazes de aprender a partir de dados, identificar padrões e realizar previsões ou tomar decisões com base nesses padrões. Essa abordagem tem ganhado destaque nos últimos anos devido à crescente disponibilidade de dados e ao avanço das capacidades computacionais.

Para Luger (2013), a Inteligência Artificial (IA) é “*o ramo da ciência da computação que se ocupa da automação do comportamento inteligente*”. Neste sentido, a IA tem sido amplamente utilizada em diversas áreas, incluindo na contabilidade e na administração tributária. Com a utilização de técnicas de IA, é possível analisar grandes quantidades de dados de forma eficiente e identificar padrões e *insights* que seriam difíceis ou impossíveis de serem detectados por um ser humano. Isso pode contribuir para a melhoria do processo de tomada de decisão na administração tributária.

Na seara das Ciências Contábeis, Silva, Costa e Pimenta (2022) explicam que a inteligência artificial tem permitido uma maior automatização de tarefas na contabilidade, como o reconhecimento automático de dados, escrituração de documentos e geração de relatórios financeiros, extração de partes relevantes, economizando tempo para contadores, permitindo que esses profissionais se dediquem a tarefas mais complexas e de tomada de decisão.

Stancheva-Todorova (2018) corrobora esse entendimento e afirma que a IA é uma oportunidade para a renovação da profissão contábil, permitindo que os contadores se concentrem em tarefas mais valiosas, embora ainda existam desafios a serem enfrentados, como atualizar os currículos de contabilidade e considerar o impacto da IA nas normas contábeis e na transparência dos dados.

Nesse contexto, considerando que a cobrança da dívida ativa tem sido um dos principais desafios da Administração Tributária, visto que muitas vezes os devedores se recusam a pagar seus débitos, a IA surge como uma alternativa promissora para melhorar o desempenho da cobrança da dívida ativa, aumentando a eficiência e a efetividade das ações de cobrança, e conseqüentemente, sua respectiva evidenciação contábil a partir de bases mais confiáveis.

A Procuradoria-Geral da Fazenda Nacional (PGFN), por exemplo, utiliza IA, a partir da técnica de *Business Intelligence* (BI) para apresentar informações relevantes, inclusive de forma gráfica, para a tomada de decisão em relação à recuperação do crédito registrado em dívida ativa da União (CRUZ, 2020).

Segundo Mitchell (1997), a aprendizagem de máquina é definida como "*o estudo de algoritmos computacionais que melhoram automaticamente através da experiência*". Isso significa que os algoritmos de aprendizagem de máquina são projetados para aprender e se adaptar com base em exemplos e dados fornecidos, sem a necessidade de programação explícita.

Um dos principais objetivos da aprendizagem de máquina é criar modelos capazes de generalizar a partir dos dados de treinamento, ou seja, serem capazes de fazer previsões precisas em novos dados que não foram vistos durante o treinamento. Isso é alcançado através de técnicas como o aprendizado supervisionado, no qual o algoritmo é treinado com exemplos rotulados, e o aprendizado não supervisionado, no qual o algoritmo busca encontrar padrões ou estruturas nos dados sem o auxílio de rótulos.

Alguns algoritmos comumente utilizados na aprendizagem de máquina incluem as árvores de decisão, as redes neurais artificiais, as máquinas de vetores de suporte e os algoritmos de agrupamento, como o *k-means*. Cada algoritmo possui suas características e propriedades específicas, sendo adequado para diferentes tipos de problemas e conjuntos de dados. Diversos pesquisadores têm se dedicado ao assunto, desenvolvendo trabalhos.

O estudo de Sá (2022) utilizou técnicas de aprendizado de máquina para identificar a propensão de fornecedores praticarem fraudes em licitações do Governo do Estado do Rio de Janeiro. Os dados foram obtidos do Portal da Transparência do governo e da Receita Federal. O algoritmo de classificação utilizado, *Support Vector Machine* (SVM), obteve um resultado promissor, com uma precisão de 99,30% e apenas um falso positivo na base de teste.

O trabalho de Santos (2020) propõe uma metodologia de previsão de tendências no mercado de ações brasileiro usando algoritmos de *machine learning*, como *Random Forest*, Redes Neurais Artificiais e *Support Vector Machines*. Indicadores técnicos são utilizados como entradas para treinar os modelos, e os resultados são avaliados por indicadores de performance.

Já pesquisa promovida por Basak et al. (2020) aborda a previsão da direção dos preços das ações das empresas *Apple* e *Facebook* utilizando a biblioteca XGBoost. Esse algoritmo é uma implementação otimizada do *gradient boosting*, que combina várias árvores de decisão em um modelo conjunto. Neste estudo, os autores utilizaram o XGBoost para construir um modelo de classificação capaz de determinar se os preços das ações irão aumentar ou diminuir com base

em dados históricos. Os resultados do estudo mostraram que o modelo desenvolvido apresentou uma alta precisão na previsão da direção dos preços das ações, superando outras abordagens de previsão existentes.

Portanto, a aprendizagem de máquina tem aplicações em diversas áreas, como reconhecimento de padrões, processamento de linguagem natural, visão computacional, sistemas de recomendação e análise de dados, entre outros. Seu potencial para impulsionar avanços tecnológicos e resolver problemas complexos tem despertado o interesse de pesquisadores e profissionais em todo o mundo.

Neste cenário, a governança de dados mostra-se relevante, visto que a gestão de dados governamentais é fundamental para o bom funcionamento da máquina pública e para o atendimento aos interesses da sociedade.

A governança de dados é um conjunto de práticas e processos que visam garantir a qualidade, a integridade e a segurança dos dados, além de promover o seu uso eficiente e estratégico. Barbieri (2020) a define como um “*conjunto de práticas, dispostas em um framework, com o objetivo de organizar o uso e o controle adequado dos dados como um ativo organizacional*”.

A adoção de uma boa governança de dados na administração pública pode trazer inúmeros benefícios, como o aumento da transparência e da eficiência na prestação de serviços públicos, a melhoria da tomada de decisão e a promoção da inovação. Por outro lado, a falta de uma governança de dados adequada pode levar a problemas como a duplicidade de informações, a perda de dados, a falta de segurança e a tomada de decisões equivocadas.

Portanto, a governança de dados é um tema crucial na administração pública, sendo fundamental para garantir a qualidade, integridade e segurança dos dados, além de promover o seu uso eficiente e estratégico.

É importante que as organizações governamentais incorporem a governança de dados em suas estratégias e promovam a colaboração e participação da sociedade na definição de políticas e práticas de gestão de dados. O desenvolvimento de soluções tecnológicas que permitam o compartilhamento seguro e eficiente de informações entre os órgãos governamentais e a sociedade também é fundamental para a promoção da governança de dados na administração pública.

1.5 Estrutura da Procuradoria Geral do Estado do Rio de Janeiro

A Procuradoria Geral do Estado do Rio de Janeiro (PGE/RJ) desempenha um papel fundamental na representação judicial e na consultoria jurídica do Estado do Rio de Janeiro. Suas atribuições estão previstas tanto na Constituição Federal, no artigo 132, quanto na Constituição do Estado do Rio de Janeiro, no artigo 176.

Como órgão central do Sistema Jurídico Estadual, a PGE/RJ supervisiona os serviços jurídicos das administrações direta e indireta do Estado. Isso significa que ela é responsável por orientar e prestar assessoria jurídica aos órgãos e entidades estaduais, garantindo que suas ações estejam de acordo com a legalidade.

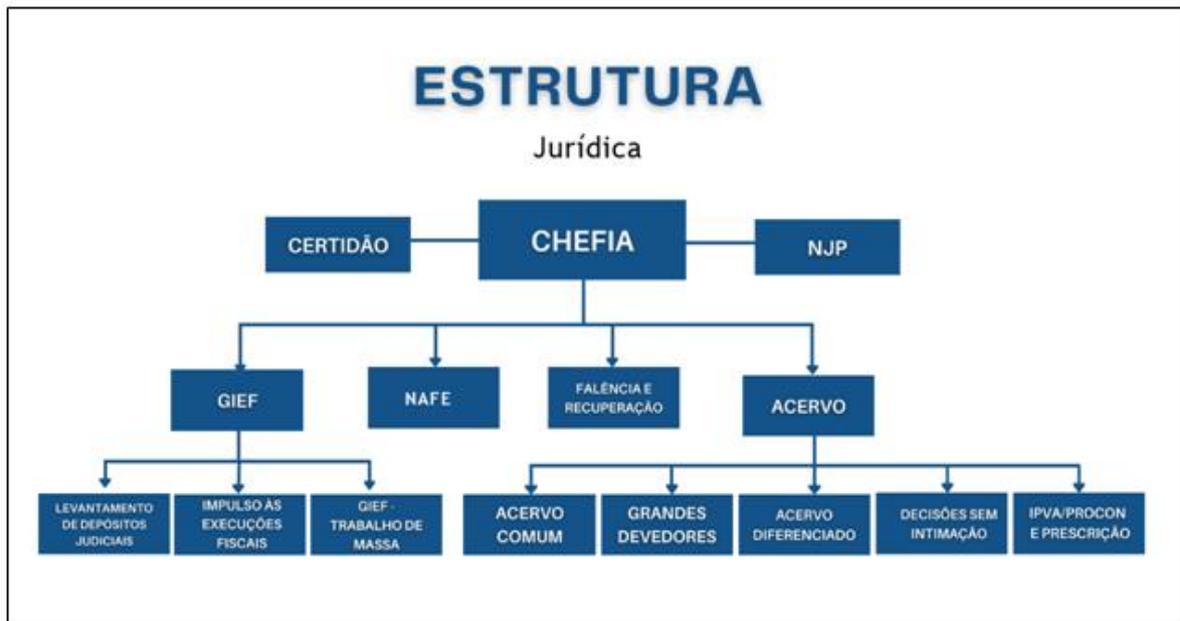
Além disso, a PGE/RJ atua no controle interno da legalidade dos atos da Administração Pública. Isso significa que ela verifica se os atos praticados pelos órgãos e entidades estaduais estão em conformidade com a legislação vigente. Essa função é importante para assegurar que o Estado atue dentro dos limites legais e para evitar possíveis irregularidades ou ilegalidades.

Outra responsabilidade da PGE/RJ é a defesa dos interesses legítimos do Estado do Rio de Janeiro, tanto na esfera judicial quanto na extrajudicial. Isso inclui a representação do Estado em processos judiciais em que ele seja parte, bem como a negociação de acordos e a condução de medidas administrativas para proteger os interesses estaduais.

A organização e o funcionamento da Procuradoria Geral do Estado do Rio de Janeiro (PGE/RJ) estão estabelecidos na Lei Complementar Estadual nº 15, de 25 de novembro de 1980. Essa lei trata não apenas da estrutura e do funcionamento da PGE/RJ, mas também da carreira de Procurador do Estado, estabelecendo as regras para o ingresso na carreira, bem como os cargos, direitos, deveres, garantias e prerrogativas dos procuradores estaduais.

A Procuradoria da Dívida Ativa está organizada em diferentes setores Jurídicos e Administrativos que atuam em áreas específicas, cujo objetivo é buscar uma administração mais gerencial, porém, sempre respeitando os regramentos legais, conforme organograma ilustrado nas Figuras 1 e 2.

Figura 1: Organograma da estrutura Jurídica da Dívida Ativa PGE/RJ



Fonte: PGE/RJ, 2023.

Figura 2: Organograma da estrutura administrativa da Dívida Ativa da PGE/RJ



Fonte: PGE/RJ, 2023.

Os setores jurídicos contemplam a Chefia de Gabinete, o Setor de Certidão e Suspensão de Exigibilidade de Créditos, o NJP - Negociação Jurídica Processual, o GIEF - Grupo de Impulso à Execução Fiscal e Levantamento de Depósitos Judiciais, o NAFE - Núcleo de Ações

Fiscais Estratégicas, o Núcleo de Falências e Recuperação Judicial, os Acervos Comum, Grandes Devedores, Diferenciado e Decisões sem Intimação, além do Núcleo IPVA/PROCON, Pagamento e Prescrição – NIPPP. O Quadro 1 apresenta a descrição das atividades dos setores jurídicos:

Quadro 1: Estrutura Jurídica da Dívida Ativa da PGE/RJ

Setor	Descrição
Certidão - Setor de Certidão e Suspensão de Exigibilidade de Créditos	Conferência jurídica das certidões de regularidade fiscal, suspensão da exigibilidade de créditos, suspensão para fins de expedição de certidão positiva com efeitos de negativa, além de outras anotações relevantes.
NJP - Negociação Jurídica Processual	Negociação direta com os contribuintes, estipulação de mudanças no procedimento ou disposição sobre ônus, poderes, faculdades e deveres processuais no âmbito das ações judiciais e execuções fiscais, visando à celeridade e economia processual.
GIEF - Grupo de Impulso à Execução Fiscal e Levantamento de Depósitos Judiciais	Prática dos atos de impulso às execuções fiscais não embargadas, investigação patrimonial e societária dos devedores, ofício nos processos de leilão judicial e atuação nas cartas precatórias. Atua em processos de apropriação de depósitos judiciais, após o levantamento dos valores nos autos e a confirmação de ingresso em receita pela SEFAZ.
NAFE - Núcleo de Ações Fiscais Estratégicas	Investigação patrimonial e societária mais profunda e uma estratégia específica de atuação, em geral casos de fraude estruturada. Além disso, está sempre em contato com o Ministério Público para atuação conjunta e com órgãos de inteligência.
Núcleo de Falências e Recuperação Judicial	Acompanhamento dos processos de recuperação judicial e falência de sociedades devedoras do Estado do Rio de Janeiro, zelo pelos seus interesses, ofício nas execuções fiscais ajuizadas em face de tais devedores, análise da viabilidade de prosseguimento da cobrança, além de acompanhamento dos precedentes judiciais na matéria e devedores estratégicos.
Acervos Comum	Atuação nos embargos à execução e seus recursos, interpôs e atuou em recursos originados em EFs e responde às exceções de pré-executividade, excetuando processos de atribuição do NIPPP.
Grandes Devedores	Atuação nas execuções fiscais envolvendo ICMS, embargadas ou não, que apresentem elevado potencial de recuperação de crédito do valor devido, a depender do faturamento e da saúde financeira da sociedade devedora - CDA com valor acima de R\$ 5 MM e inscritas a partir de 2009.
Acervo Diferenciado	Atuação nas execuções fiscais embargadas ou não, cujo valor de face das CDA esteja entre R\$ 2 e 5 MM e acima de R\$ 5 MM se inscritas até 2008

Fonte: O autor, 2023.

Já entre os setores administrativos engloba o Setor de Triagem e Abertura de PA, o Setor de Tramitação de PA - STPA, o Setor de Ofícios e Classificação e o Núcleo de Pesquisa - NUPED. Cada um desses setores é responsável por diferentes tarefas, como por exemplo, a conferência de certidões fiscais, a negociação com contribuintes, o impulso às execuções fiscais, a investigação patrimonial e societária, a atuação em processos de recuperação judicial e falência, a análise de processos judiciais sem intimação, entre outras. O Quadro 2 apresenta a estrutura administrativa da Procuradoria:

Quadro 2: Estrutura administrativa da Dívida Ativa da PGE/RJ

Setor	Descrição
Triagem e Abertura de PA	Classifica e distribui os processos judiciais conforme o valor e a matéria no PGE Digital e inaugura os processos administrativos para acompanhamento de processos judiciais, quando necessário.
Tramitação de PA - STPA	Recebe todos os processos administrativos e ofícios que ingressam na Procuradoria da Dívida Ativa, pelo Sistema SEI. Os processos são direcionados ao setor competente e os ofícios, ao Procurador do feito. Atua também como apoio aos Procuradores e analistas.
Ofícios e Classificação	Classifica e distribui os processos judiciais conforme o valor e a matéria no PGE Digital e inaugura os processos administrativos para acompanhamento de processos judiciais, quando necessário.
Núcleo de Pesquisa - NUPED	Responsável pelo levantamento de dados patrimoniais e cadastrais do devedor em mais de 20 bases de dados disponíveis, com destaque para Sistema CPF/CNPJ/SERPRO, JUCERJA, RCPJ, CAGED e Sistema de Consultas a Autos de Infração da SEFAZ-RJ.
DAC - Departamento de Anotação e Inscrição em Dívida Ativa e Expedição de Certidão	Inscreve as notas de débitos recebidas dos órgãos da Administração que ainda não enviam de forma automatizada, realiza todas as anotações e alterações/retificações dos dados relativos às CDA no Sistema da Dívida Ativa, mediante determinação/autorização da chefia da PG05. Atua também na expedição de Certidão Positiva e Certidão Positiva com Efeito de Negativa (CPEN).
SPRO - Setor de Protocolo	Atendimento ao público virtual. Orienta e inaugura processos que tratam de pedidos de revisão do cálculo e/ou cancelamento da CDA, responsável por dar andamento ao pedido de Certidão de Regularidade Fiscal e pela abertura de requerimentos administrativos diversos no SEI, bem como pela intimação do contribuinte.
SPCA - Setor de Parcelamento e Cobrança Administrativa	Atendimento ao público, virtual e presencial, realiza simulação e processa os pedidos de parcelamento; emissão de DARJ; cadastramento e acompanhamento do parcelamento deferido. Pesquisa de débitos; Atendimento às Procuradorias Regionais; Cancelamento de débito; Cálculo de saldo remanescente; apostilamento e Atualização da certidão.
TI/DA	Suporte técnico, assim como o de desenvolvimento dos sistemas para auxiliar no funcionamento da Especializada. Implementação e atuação nos projetos tecnológicos de toda a especializada.
Agente de Pessoal	Controla todas as demandas de pessoal da especializada (Procuradores, Servidores, Residentes Jurídicos e Estagiários), tais como, o ingresso e desligamento de funcionários, gerir as folhas de ponto, controle de férias e substituições.
Núcleo de Gestão	Atua na formulação e mensuração de indicadores de desempenho e no mapeamento de processos na especializada. Elaboração de relatórios gerenciais com vistas a tomadas de decisão pela chefia. Otimização e ampliação dos métodos de cobrança administrativa dos créditos (Cobrança amigável, Protesto de Títulos etc.). Envio de relatórios com informações das CDA para outros órgãos como TJ e TCE. Mapeamento de processos, aperfeiçoamento das rotinas, elaboração de fluxos e roteiros para as novas rotinas da Especializada. Suporte técnico Contábil nas demandas fiscais, revisão de cálculos, evolução do estoque e da recuperação da dívida ativa; dar suporte na prestação de contas da dívida ativa para TCE e demais órgãos governamentais.

Fonte: O autor, 2023.

Cabe registrar que as informações sobre a dívida ativa são compartilhadas com partes interessadas. Assim, a PGE/RJ (Procuradoria-Geral do Estado do Rio de Janeiro) encaminha tais dados para os órgãos oficiais do Estado que têm interesse direto nessa informação, como a Secretaria de Fazenda do Estado do Rio de Janeiro (SEFAZ/RJ), o Tribunal de Contas do Estado

do Rio de Janeiro (TCE/RJ), a Controladoria Geral do Estado do Rio de Janeiro (CGE/RJ) e o Rioprevidência¹.

1.6 Análise da metodologia atual adotada para ajuste de perdas para dívida ativa no Estado do Rio de Janeiro

Da análise da metodologia atual (Anexo A), é possível verificar que considera a segmentação dos créditos não tributários inscritos antes e após 1997 e atribui faixas de percentual de dificuldade de acordo com a descrição das dificuldades (baixa, média, alta) e suas respectivas faixas de percentual (0,00 a 25,00%, 26,00 a 50,00%, 51,00 a 100,00%). Além disso, considera outros aspectos relevantes a serem considerados na análise de créditos da Dívida Ativa, que incluem:

1. Não ajuizado, valor menor: Refere-se aos créditos que ainda não foram levados a juízo, mas possuem valores menores em comparação aos demais.
2. Com exigibilidade suspensa: São créditos em que a exigibilidade do pagamento está suspensa devido a algum motivo específico, como uma decisão judicial ou outros fatores legais.
3. CDA de empresas com registro de falência/concordata: São créditos de empresas que estão em processo de falência ou concordata, o que pode afetar a possibilidade de recuperação desses créditos.
4. Com parcelamento interrompido (PDA ou SEF): Refere-se aos créditos em que o parcelamento foi interrompido, seja por falta de pagamento das parcelas ou por outras razões relacionadas ao Programa de Parcelamento ou ao Sistema Especial de Regularização Tributária.
5. CDA com exigibilidade plena amigável: São créditos em que a exigibilidade do pagamento é plena, mas existe a possibilidade de negociação amigável para a quitação do débito.

¹ Para atendimento do disposto na Lei 3.189/1999.

6. CDA com exigibilidade plena judicial: São créditos em que a exigibilidade do pagamento é plena e o processo está em fase judicial, ou seja, já foram ajuizados e estão seguindo os trâmites legais.

Portanto, a presente subseção atende ao **objetivo de pesquisa 2, que é "analisar a metodologia atual, adaptada pela PGE/RJ, para ajuste dos valores das perdas"**.

2 METODOLOGIA

Nesta seção serão descritos o método e os passos utilizados para a elaboração da proposta de pesquisa, iniciando pela população e amostra, seguida da coleta e tratamento dos dados, que terá como produto final a construção da base de dados para o desenvolvimento do *rating* (levantamento de fatores) dos créditos da dívida ativa.

Para Fachin (2001) método é “um instrumento de conhecimento que proporciona aos pesquisadores, orientação para facilitar e planejar uma pesquisa, formular hipóteses, coordenar investigações, realizar experiências e interpretar resultados”. Hammersley e Atkinson (1995, p. 24) afirmam que “o desenho de pesquisa deve ser um processo reflexivo que opera em cada etapa de um projeto”.

A presente pesquisa utilizará o método dedutivo. Na concepção de Matias-Pereira (2016, p.50) o raciocínio dedutivo tem o objetivo de explicar o conteúdo das premissas. Quanto aos meios técnicos para investigação, será utilizado o método estatístico, que segundo o mesmo autor, trabalha com níveis de confiança de que os dados correspondem aos fatos.

Ademais, a presente pesquisa será quantitativa, pois emvidou esforços para estudar as sequências numéricas utilizadas para representar os fenômenos investigados.

2.1 Identificação das bases de dados

Para viabilizar o estudo, o primeiro passo consistiu na identificação das bases de dados existentes que poderiam contribuir com o constructo do *rating*. A partir daí, verificou-se como passíveis de utilização o Sistema da Dívida Ativa (SDA) da PGE-RJ, os dados da Receita Federal contendo o Cadastro Nacional de Pessoas Jurídicas (CNPJ), bem como da base da PGFN que demonstra as empresas com dívidas inscritas na União.

A busca por outras bases, além daquelas institucionais da PGE mostrou-se relevante na medida em que estes dados permitirão análises mais aprofundadas. Agresti e Finlay (2012) definem que “a coleta de informação é o coração de todas as ciências, é o que fornece as observações usadas na análise estatística. As observações coletadas sobre as características de interesse são chamadas coletivamente de dados.”

Com o intuito de avaliar o comportamento mais recente dos devedores, eliminando o efeito de condutas que podem ter sido extintas ao longo do tempo por alterações na legislação

ou nos processos, optou-se por empregar na pesquisa apenas os dados mais atuais da dívida ativa, referentes aos últimos cinco anos (2018-2022).

2.2 Variáveis da pesquisa

Após a identificação das bases de dados, foi selecionada como variável dependente a recuperabilidade do crédito de dívida ativa. Ela indica se o crédito foi recuperado (1) ou não (0).

Na sequência, com o objetivo de identificar padrões de recuperabilidade do crédito foram selecionadas 10 (dez) variáveis para a criação das faixas de classificação (*rating*): valor total, ajuizada, exigibilidade suspensa, situação do parcelamento, quantidade de dias, região, situação cadastral, dias de empresa, situação na PGFN e setor econômico. Através da análise dessas variáveis, busca-se compreender quais características e informações estão associadas a uma maior ou menor probabilidade de recuperação dos créditos de dívida ativa. É esperado comportamento positivo de todas as variáveis estudadas.

O “valor total” é uma variável explicativa que representa o valor atualizado da dívida em reais até 2022. Essa variável fornece informações sobre a magnitude da dívida e pode influenciar a probabilidade de recuperação do crédito.

A variável “ajuizada” indica se o crédito está sujeito a processo judicial ou não. Ela é uma variável binária, assumindo o valor 1 quando o crédito está sujeito a processo judicial e 0 quando não está. Essa variável pode afetar a probabilidade de recuperação do crédito, uma vez que ações judiciais podem impactar o tempo e os recursos necessários para a recuperação.

A variável “exigibilidade suspensa” indica se a cobrança do tributo está suspensa ou não. Quando a variável assumir o valor 1, significa que a cobrança do tributo está suspensa de acordo com as hipóteses do artigo 151 do CTN (1) ou não (0). Essa variável pode influenciar a probabilidade de recuperação do crédito, pois uma suspensão da cobrança pode afetar a capacidade de recuperação.

A situação do parcelamento é outra variável explicativa que indica se as Certidões de Dívida Ativa estão parceladas ou não. Essa variável assume o valor 1 quando as CDA estão parceladas e 0 quando não estão. O parcelamento pode ter impacto na capacidade de recuperação do crédito, pois afeta os termos e condições de pagamento.

A quantidade de dias é uma variável explicativa que mede o período de tempo entre a inscrição do tributo em dívida ativa e a recuperação do crédito (caso recuperado) ou até

31/12/2022 (caso não recuperado). Essa variável pode indicar a eficiência e o tempo necessário para a recuperação do crédito.

A variável região divide os créditos inscritos em três macrorregiões: região metropolitana do RJ, interior do RJ e outros estados. Essa variável geográfica pode influenciar a probabilidade de recuperação do crédito, pois diferentes regiões podem apresentar características econômicas e jurídicas distintas.

A situação cadastral é uma variável explicativa que indica a situação cadastral das empresas com créditos em dívida ativa, com base nas informações do CNPJ. Ela abrange diferentes categorias, como nula, ativa, suspensa, inapta, baixada e não encontrada. Essa variável pode refletir a condição legal e operacional das empresas, o que pode afetar a recuperação do crédito.

A situação "Nula" indica a ausência de registros com essa classificação, ou seja, nenhum CNPJ é considerado nulo nas CDA analisadas. Por outro lado, a situação "Ativa" representa o CNPJ que está em plena conformidade com suas obrigações fiscais e possuem uma situação regular. A situação "Suspensa" indica que alguns CNPJ têm sua situação cadastral suspensa devido a processos administrativos ou judiciais em andamento que afetam a regularidade da empresa. Já a situação "Inapta" é atribuída aos CNPJ que não cumprem determinadas obrigações legais, como a entrega de declarações fiscais ou o pagamento de tributos, resultando em sua classificação como inaptos. Os CNPJ classificados como "Baixada" estão encerrados oficialmente ou tiveram sua atividade econômica encerrada. Por fim, a situação "não encontradas" indica que alguns CNPJ não foram encontrados nas bases de dados analisadas.

A variável explicativa "Dias de empresa" calcula a quantidade de dias desde a abertura da empresa até 31/12/2022, tanto para empresas com créditos recuperados quanto para empresas com créditos não recuperados. Essa variável pode indicar a experiência e a estabilidade das empresas, o que pode influenciar a recuperação do crédito.

A situação na PGFN é uma variável explicativa que verifica se a empresa possui dívidas inscritas na União. Essa variável assume o valor 1 quando a empresa possui dívidas inscritas na União e 0 quando não possui. A presença de dívidas inscritas na União pode ter impacto na recuperação do crédito, uma vez que pode influenciar as ações e medidas tomadas pela instituição responsável pela cobrança.

A variável "setor econômico" demonstra em qual setor produtivo da economia a empresa está inserida, considerando os 3 principais grupos, setor primário, setor secundário e setor terciário.

O setor primário engloba as atividades econômicas relacionadas à extração de recursos naturais, como agricultura, pecuária, pesca, mineração e exploração florestal. É o setor responsável pela produção de matérias-primas.

O setor secundário, por sua vez, é o setor industrial. Compreende as atividades de transformação das matérias-primas em produtos acabados ou semiacabados. Inclui indústrias de manufatura, construção civil, energia, produção de bens de consumo duráveis e não duráveis, entre outros.

O setor terciário é o setor de serviços, que abrange uma ampla gama de atividades voltadas para atender às necessidades da população e das empresas. Inclui serviços de comércio, transporte, turismo, educação, saúde, finanças, comunicações, tecnologia da informação, entretenimento, entre outros. O setor terciário é caracterizado pela prestação de serviços em vez da produção de bens físicos.

A Tabela 1 reúne as mencionadas variáveis explicativas:

Tabela 1: Variáveis explicativas do estudo

Variável	Descrição da Variável	Descrição da variável
V ₁	Valor total	Representa o valor atualizado da dívida em reais até 2022.
V ₂	Ajuizada	Demonstra se o crédito está sujeito a processo judicial (1) ou não (0).
V ₃	Exigibilidade suspensa	Indica se a cobrança do tributo está suspensa (1) ou não (0).
V ₄	Situação do parcelamento	Indica se as Certidões de Dívida Ativa (CDA) estão parceladas (1) ou não (0).
V ₅	Quantidade de dias	Refere-se ao período em dias entre a inscrição do tributo em Dívida ativa e a recuperação do crédito (caso recuperado) ou até 31/12/2022 (caso não recuperado).
V ₆	Região	Divide os créditos inscritos em três macrorregiões: 1 (região metropolitana do RJ), 2 (interior do RJ) e 3 (outros estados).
V ₇	Situação cadastral	Indica a situação cadastral das empresas com créditos em Dívida Ativa com base nas informações do CNPJ com base nas informações do CNPJ: 1 (nula), 2 (ativa), 3 (suspensa), 4 (inapta), 5 (baixada) e 6 (não encontrada).
V ₈	Dias de empresa	Calcula a quantidade de dias desde a abertura da empresa até 31/12/2022 para empresas com créditos recuperados ou não recuperados.
V ₉	Situação na PGFN	Verifica se a empresa possui dívidas inscritas na União (1) ou não (0).
V ₁₀	Setor econômico	Identifica em qual dos grandes grupos da economia a empresa está inserida: 1 (setor primário), 2 (setor secundário) e 3 (setor terciário).

Fonte: O autor, 2023.

2.3 Coleta e preparação dos dados

Os dados utilizados estão distribuídos em três bases de dados, nomeadamente o Sistema da Dívida Ativa (SDA) da PGE-RJ, o Cadastro Nacional de Pessoas Jurídicas (CNPJ) da Receita Federal e a base da PGFN que demonstra as empresas com dívidas inscritas na União. Para se ter uma ideia, a base de dados do CNPJ possui mais de 55 milhões de linhas.

A variável usada na junção (*merge*) das três bases de dados foi o CNPJ, presente em todos os três grupos de informação utilizados. A junção das bases foi realizada por meio de comandos de linguagem de programação SQL (*Structured Query Language*).

Após o processo de junção dos dados das bases de dados, foi empregada a ferramenta *Python*, muito utilizada na programação, por ser uma linguagem simples, versátil e com ampla gama de ferramentas lógicas para os mais variados trabalhos. É uma linguagem de programação interpretada, de alto nível e orientada a objetos, que é uma das mais populares do mundo. Ela é usada em diversas aplicações, desde a geração de relatórios, ciência de dados e aprendizado de máquina até desenvolvimento web e automação de tarefas.

Ademais, foi realizado um tratamento prévio dos dados, com vistas a identificar eventuais inconsistências. As lacunas de dados, dados fora do intervalo ou inconsistentes logicamente foram excluídos caso a caso.

2.4 Análise dos dados

Inicialmente, foi feita uma tentativa de utilizar regressão logística para resolver o problema de classificação da recuperabilidade do crédito de dívida ativa. No entanto, após a análise dos resultados obtidos, foi constatado que o modelo foi um classificador fraco, motivo pelo qual não se mostrou condizente com os objetivos da pesquisa. Diante disso, foi necessário buscar alternativas para alcançar uma solução mais precisa e eficaz para o problema em questão.

Um classificador fraco pode ser definido como aquele que tem uma capacidade de generalização apenas ligeiramente melhor do que a escolha aleatória. Em contrapartida, um classificador forte é capaz de generalizar bem e ter uma precisão alta. Schapire (1990) apresentou uma técnica conhecida como amplificação de classificadores fracos, que consiste em combinar vários classificadores fracos para formar um classificador forte e obter uma precisão mais alta. Essa técnica tem sido amplamente utilizada em aprendizado de máquina e foi um marco no desenvolvimento de algoritmos de aprendizado de máquina baseados em *ensemble*.

Ensemble é uma técnica utilizada em aprendizado de máquina para combinar vários modelos de predição individuais em um modelo maior e mais poderoso. Esses modelos individuais podem ser de diferentes tipos, como árvores de decisão, redes neurais artificiais, regressão, entre outros. O objetivo do *ensemble* é criar um modelo mais preciso do que qualquer um dos modelos individuais usados para criá-lo.

Existem várias técnicas de *ensemble*, dentre as mais usuais destaca-se o *boosting*, em que modelos mais fracos são iterativamente treinados e combinados para formar um modelo mais forte.

Foi então que se optou pelo uso do algoritmo de *boosting* da biblioteca XGBoost (*Extreme Gradient Boosting*) em Python, que permitiu uma abordagem mais robusta e avançada para o treinamento do modelo de classificação. Além disso, foram utilizadas outras bibliotecas de suporte para a análise e visualização de dados, como o Pandas e o *Matplotlib*.

O XGBoost é um modelo de aprendizado de máquina do tipo *boosting* altamente eficiente e originado do modelo de árvore de decisão. Ele usa a classificação de árvore para obter resultados de previsão melhores e com maior eficiência operacional.

Uma árvore de decisão é uma estrutura de dados em que cada nó interno representa um teste em um atributo, cada ramo representa o resultado do teste e cada folha representa uma classe ou valor de saída.

Portanto, ele utiliza árvores de decisão como modelos básicos e, durante o processo de treinamento, aprimora iterativamente essas árvores adicionando novas árvores que corrigem os erros dos modelos anteriores. Isso resulta em um modelo final que é uma combinação ponderada das previsões de várias árvores. Harrison (2020, p. 128) explica que “ele cria uma árvore fraca e, então “melhora” as árvores subsequentes (faz um *boosting*) a fim de reduzir os erros residuais. O algoritmo tenta capturar e tratar qualquer padrão nos erros, até que pareçam ser aleatórios”.

Uma árvore de classificação é um modelo que é capaz de classificar ou prever uma saída categórica a partir de um conjunto de entradas. A árvore de classificação é construída a partir de um conjunto de regras de decisão que são aprendidas a partir dos dados de treinamento (BREIMAN et al., 1984).

O XGBoost tem sido amplamente utilizado em muitas áreas, como previsão de falência bancária, negociação financeira e detecção de intrusão de rede. Estudos recentes compararam o desempenho do XGBoost com outros modelos, como regressão logística, árvore de decisão e redes neurais artificiais, e verificaram sua superioridade em avaliação (LI et al., 2019).

Além disso, o XGBoost possui uma série de recursos que o tornam uma escolha popular para problemas de aprendizado de máquina, incluindo a capacidade de lidar com dados ausentes, regularização para evitar *overfitting*, processamento paralelo e suporte a múltiplas funções de perda (CHEN e GUESTRIN, 2016; FRIEDMAN, 2001).

No caso específico em que foi utilizado para resolver um problema de classificação de recuperabilidade de crédito de dívida ativa, o XGBoost mostrou-se uma escolha acertada, uma vez que permitiu obter uma maior acurácia na classificação das amostras.

Portanto, o modelo descrito foi baseado em árvores de decisão que foi treinado para resolver um problema de classificação, onde a variável-alvo é a recuperabilidade do crédito de dívida ativa. Foram consideradas as dez variáveis para treinar o modelo, conforme detalhado na subseção 3.2.

O treinamento do modelo envolveu a definição dos hiperparâmetros do modelo, que para o caso deste trabalho, foi de 200 para o número de árvores (*n_estimators*), 0.1 para o “*base_score*” que é o valor inicial atribuído a todas as instâncias de dados antes de iniciar o treinamento do XGBoost e é comumente usado quando se trata de problemas de regressão.

O *max_depth* ficou definido no nível 7 e se refere a profundidade máxima de cada árvore de decisão, por fim, o “*min_child_weight*” foi fixado no nível 6, e diz respeito a definição do peso (instâncias de treinamento) necessária em um nó folha para continuar dividindo-o em subárvores adicionais. O pseudocódigo do XGBoost é descrito a seguir:

Algoritmo 1 - Pseudocódigo

```

Definir dados de treinamento  $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$ 
Definir o número de árvores de decisão  $T$ 
Definir a taxa de aprendizagem  $\eta$ 
Definir a profundidade máxima da árvore  $max\_depth$ 
Definir o valor do parâmetro de regularização  $\lambda$ 
Definir o valor do parâmetro de controle de overfitting  $\gamma$ 
Definir o número mínimo de amostras em uma folha  $min\_child\_weight$ 
Definir o número de threads a serem usados na computação paralela  $n\_threads$ 
Definir o número de classes  $K$  (para problemas de classificação)
Para  $k$  de 1 a  $K$  (se for um problema de classificação):
    Definir dados de treinamento  $D_k = \{(x_1, y_{1_k}), (x_2, y_{2_k}), \dots, (x_m, y_{m_k})\}$ 
    Definir os resíduos iniciais  $r = \{0, 0, \dots, 0\}$ 
Para  $t$  de 1 a  $T$ :
    Para  $k$  de 1 a  $K$  (se for um problema de classificação):
        Calcular os gradientes e os Hessians para cada amostra de treinamento:
             $grad, hess = gradient\_hessian(y_k, y\_pred_k)$ 
        Definir os resíduos iniciais  $r = -grad$ 
    Para  $j$  de 1 a  $J$  (o número de nós da árvore de decisão):
        Calcular os gradientes e os Hessians para cada amostra de treinamento:
             $grad, hess = gradient\_hessian(y, y\_pred)$ 
        Para cada nó  $j$ , encontrar o melhor split que minimiza a função de perda:
             $split = melhor\_split(j, grad, hess)$ 
        Se o ganho de informação do split for menor do que zero:
            Parar a construção da árvore para esse nó
        Se não:
            Adicionar os filhos do nó  $j$  à árvore
            Definir os resíduos  $r$  para cada amostra de treinamento em cada filho:
                 $r\_left = -grad[x]$  se  $x$  está à esquerda do split, 0 caso contrário
                 $r\_right = -grad[x]$  se  $x$  está à direita do split, 0 caso contrário
    Para  $k$  de 1 a  $K$  (se for um problema de classificação):

```

Para cada amostra de treinamento xi:
 Fazer a predição para cada amostra de treinamento xi e atualizar a previsão do modelo: $y_pred_k(xi)$
 $= y_pred_k(xi) + \eta * \text{predição da árvore para a classe } k$
 Calcular os resíduos atualizados para cada amostra de treinamento $r_k = y_k - y_pred_k$
 Fazer a previsão final para cada amostra de treinamento xi somando as previsões de todas as árvores e classes. Se o ganho de informação da última árvore for menor do que zero: Parar o treinamento

Fonte: O autor, 2023.

A avaliação da qualidade de um modelo é uma etapa crucial para garantir sua eficiência e confiabilidade. Para esse fim, serão empregadas métricas que permitem medir e analisar seu desempenho. Para tanto serão utilizadas as seguintes métricas: matriz de confusão, precisão (*precision*), revocação (*recall*), *f-1 score*, *support*, acurácia (*accuracy*), média macro (*macro avg*), média ponderada (*weighted avg*), ROC e AUC.

A avaliação do desempenho de um classificador pode ser feita por meio do uso de uma matriz de confusão. Essa matriz fornece uma quantificação das frequências das predições corretas e incorretas em cada classe. Em um determinado conjunto de dados, as linhas da matriz representam as classes verdadeiras, enquanto as colunas representam as classes preditas pelo classificador (FACELI et. al., 2021, p. 150).

Segundo Malhotra (2019, p. 484) a matriz de confusão ilustrada na Tabela 2, também conhecida por matriz de classificação ou matriz de predição, é uma representação tabular que contém informações sobre os casos classificados corretamente e incorretamente. Os casos classificados corretamente são dispostos na diagonal da matriz. Os elementos fora da diagonal representam os casos que foram classificados de forma incorreta. A soma dos elementos da diagonal, dividida pelo número total de casos, fornece a proporção de acertos, também conhecida como acurácia do modelo de classificação (MONARD e BARANAUSKAS, 2003).

Tabela 2: Matriz de confusão para classificador com duas classes

Classe	Predita C ₊	Predita C ₋	Taxa de erro da classe	Taxa de erro total
Verdadeira C ₊	Verdadeiros Positivos T _P	Falsos negativos F _N	$\frac{FN}{TP + FN}$	$\frac{FP + FN}{n}$
Verdadeira C ₋	Falsos Positivos F _P	Verdadeiros negativos T _N	$\frac{FP}{FP + TN}$	

Fonte: Monard e Baranauskas, 2003.

Os valores presentes na matriz de confusão são contagens que representam a quantidade de exemplos classificados corretamente ou de forma incorreta pelo modelo. Esses valores são divididos em quatro categorias:

- Verdadeiros positivos (TP): quantidade de exemplos da classe positiva (CDA recuperada) corretamente classificados como positivos .
- Falsos positivos (FP): quantidade de exemplos da classe negativa (CDA não recuperada) incorretamente classificados como positivos.
- Verdadeiros negativos (TN): quantidade de exemplos da classe negativa (CDA não recuperada) corretamente classificados como negativos.
- Falsos negativos (FN): quantidade de exemplos da classe positiva (CDA recuperada) incorretamente classificados como negativos.

Ademais, a partir dos valores da matriz de confusão, métricas como precisão, revocação, *F1-score* e acurácia podem ser calculadas, fornecendo uma visão mais abrangente do desempenho do modelo de classificação (VENTURINI, 2021, p. 23 apud WITTEN et al., 2011, 164-177).

A precisão é a proporção de exemplos positivos corretamente classificados em relação ao total de exemplos classificados como positivos, conforme indicado a seguir:

$$precision = \text{verdadeiros positivos} / (\text{verdadeiros positivos} + \text{falsos positivos}) \quad (1)$$

Já a revocação (*recall*), também conhecida como taxa de verdadeiros positivos, mede a proporção de exemplos positivos corretamente classificados em relação ao total de exemplos realmente positivos. Em outras palavras, é a capacidade do modelo de encontrar corretamente todos os exemplos positivos. A equação para calcular o *recall* consiste em:

$$recall = \text{verdadeiros positivos} / (\text{verdadeiros positivos} + \text{falsos negativos}) \quad (2)$$

O *F1-score* é uma medida que combina a precisão e o *recall* em uma única métrica resumida. É a média harmônica da precisão e do *recall*, fornecendo uma medida equilibrada entre as duas métricas. O *F1-score* é especialmente útil quando há um desequilíbrio de classes no conjunto de dados. Quanto mais próximo de 1, melhor é o desempenho do modelo. A equação para calcular o *F1-score* é:

$$F1\text{-score} = 2 * (precision * recall) / (precision + recall) \quad (3)$$

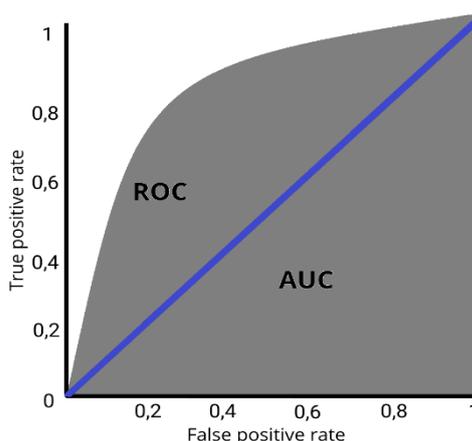
A acurácia (*accuracy*) é uma medida que calcula a proporção de exemplos corretamente classificados em relação ao total de exemplos. Em outras palavras, é a taxa de acerto global do modelo. A equação para calcular a acurácia é:

$$accuracy = (\text{verdadeiros positivos} + \text{verdadeiros negativos}) / (\text{verdadeiros positivos} + \text{verdadeiros negativos} + \text{falsos positivos} + \text{falsos negativos}) \quad (4)$$

A curva ROC (*Receiving Operating Characteristics*) ilustrada na Figura 3 é uma forma alternativa de avaliar classificadores em problemas binários, nos quais existem apenas duas classes (FACELI et. al., 2021, p. 158 apud FAWCETT, 2005).

Segundo Faceli (2021) a curva ROC é um gráfico no quais os eixos das abscissas e das ordenadas representam as taxas de falsos positivos (TFP) e taxas de verdadeiros positivos (TVP), respectivamente. O desempenho de um classificador pode ser representado por um ponto nessa curva. A interpretação da curva ROC está relacionada à sua forma e à sua proximidade com o canto superior esquerdo do gráfico, quanto mais próxima a curva estiver desse canto, melhor será o desempenho do classificador. Contudo, se a curva estiver próxima da diagonal, o modelo não possui poder discriminatório.

Figura 3: modelo de curva ROC



Fonte: O autor, 2023.

A medida AUC (*Area Under the Curve*), por seu turno, produz valores entre 0 e 1, sendo que valores mais próximos de 1 são considerados melhores. Um valor de 0,5 indica que o modelo tem um desempenho equivalente ao de um modelo aleatório, enquanto valores abaixo de 0,5

indicam um desempenho pior do que o aleatório. Ao comparar dois ou mais algoritmos utilizando essa medida, o algoritmo com AUC mais próximo de 1 é considerado superior (FACELI et. al., 2021).

2.5 Rating do potencial de recuperabilidade da dívida ativa

A variável resposta é binária (com 1 indicando CDA recuperada e 0 CDA não recuperada), enquanto a saída do XGBoost é um número no intervalo entre 0 e 1. Assim, a resposta do XGBoost foi interpretada como sendo uma estimativa da probabilidade de recuperação da CDA, condicionada aos valores das variáveis explicativas.

Visando fornecer uma classificação das CDA com base em seu potencial de recuperabilidade da dívida ativa, foi estabelecido o *rating* com as sete categorias indicadas na Tabela, cada uma abrangendo uma faixa do intervalo entre 0 e 1. O resultado é uma escala decrescente do potencial ou probabilidade de recuperação dos créditos inscritos na dívida ativa, onde o nível “A” (85% A 100%) é considerado o melhor *rating* e o nível “G” (0,01% a 9,99%) o pior.

Tabela 3: *Rating* do potencial de recuperabilidade

Nível	Resposta do XGBoost
A	85,00% A 100%
B	70,00% a 84,99%
C	55,00% a 69,99%
D	40,00% a 54,99%
E	25,00% a 39,99%
F	10,00% a 24,99%
G	0,01% a 9,99%

Fonte: O autor, 2023.

3 DESCRIÇÃO E ANÁLISE DOS RESULTADOS

Nesta seção serão detalhados os resultados obtidos na pesquisa conforme aplicação da metodologia. Primeiro será realizada uma análise descritiva dos dados e apresentada a metodologia atual adotada pela PGE/RJ e o modelo gerado pelo presente estudo para contabilização da perda.

3.1 Análise descritiva da base de dados

A base de dados inicial utilizada nesse trabalho foi a da dívida ativa do Estado do Rio de Janeiro, responsável por gerar a maior parte das informações utilizadas nesse trabalho. Os dados foram levantados para os últimos cinco anos, de 2018 a 2022, e incluíram informações sobre os valores recuperados nesse período, bem como o arquivo completo contendo todas as Certidões de Dívida Ativa (CDA) inscritas até o ano de 2022. Além da base de dados da dívida ativa, foi utilizada uma base de dados disponibilizada pelo Governo Federal, encontrada no site dados.gov, conforme detalhado no Capítulo 3, mais a base da PGFN, retirada do site da PGFN.

O Estoque da dívida ativa é subdividido em créditos tributários, tais como: ICMS, IPVA, ITBI, ITD e créditos não tributários, nos quais temos por exemplo, multas e taxas. A Tabela 4 apresenta a composição do estoque ao longo dos anos de 2018 a 2022, demonstrando o valor total, que é o somatório de créditos tributários e não tributários, o valor total dos créditos tributários e o ICMS, justamente para demonstrar a relevância desse imposto em relação ao montante total do estoque e em relação ao grupo dos créditos tributários, cujos valores estão expressos em bilhões de reais.

O estoque relacionado ao Imposto sobre Circulação de Mercadorias e Serviços (ICMS) é detalhado, evidenciando a quantidade de recursos inscritos em cada exercício.

É possível verificar um crescimento constante do estoque de ICMS ao longo dos anos, com um aumento significativo em 2022, o qual corresponde ao valor total de R\$129.974.812.130,00.

Tabela 4: Composição do estoque da dívida ativa total ao longo dos anos (2018-2022)

Estoque	2022	2021	2020	2019	2018
ICMS	129.974.811.975,31	118.086.471.012,96	114.646.369.540,42	110.654.351.114,48	103.143.306.938,77
Tributária	131.116.667.835,92	119.476.884.020,98	116.053.620.712,14	111.936.452.040,78	104.504.985.210,76
Total	135.793.207.387,76	121.977.609.267,40	118.439.186.473,93	114.296.760.757,20	106.750.103.335,50

Fonte: O autor, 2023.

A análise da Tabela 5 apresenta o crescimento nominal do estoque da dívida ativa ao longo dos anos em termos de variação percentual. Os dados evidenciam um aumento gradual tanto no estoque de ICMS quanto na dívida tributária, indicando um crescimento significativo dessas categorias ao longo do período analisado.

No caso do ICMS, é possível verificar um crescimento nominal consistente, com variações percentuais que oscilaram entre 2,99% e 10,07%. Isso demonstra um aumento expressivo no estoque de ICMS ao longo dos anos, com destaque para o ano de 2022, em que ocorreu um incremento significativo. A dívida tributária também seguiu uma tendência de crescimento nominal, com variações percentuais que variaram entre 2,95% e 9,75%.

Quanto ao estoque total da dívida ativa, que engloba a dívida tributária, e, conseqüentemente, a dívida de ICMS, a variação percentual variou entre 2,99% e 11,35%. Esses números indicam um crescimento nominal geral do estoque total ao longo do período analisado.

Tabela 5: Crescimento da dívida ativa ao longo dos anos (2018-2022)

Ano	Varição Percentual ICMS	Varição Percentual Tributária	Varição Percentual Total
2022	10,07%	9,75%	11,35%
2021	2,99%	2,95%	2,99%
2020	3,61%	3,69%	3,63%
2019	7,28%	7,09%	3,63%

Fonte: O autor, 2023.

A Tabela 6 destaca a participação do ICMS no estoque total da dívida ativa durante o período estudado. A participação do ICMS no estoque total da dívida ativa manteve-se alta e estável ao longo dos anos, variando entre 95,72% e 96,81%, indicando que o ICMS é um componente relevante na composição da dívida ativa.

Tabela 6: Participação do ICMS no estoque total da dívida ativa ao longo dos anos (2018-2022)

Estoque	2022	2021	2020	2019	2018
Participação do ICMS no estoque total da Dívida ativa	95,72%	96,81%	96,80%	96,81%	96,62%
Participação do ICMS no estoque total da Dívida ativa tributária	99,13%	98,84%	98,79%	98,85%	98,70%

Fonte: O autor, 2023.

A Tabela 7 destaca os **valores da dívida ativa recuperados durante o período de 2018 a 2022, que totalizou um montante de R\$3.757.540.294,45**. No entanto, é importante esclarecer que o valor efetivamente liquidado das Certidões de Dívida Ativa (CDA) inscritas foi de R\$2.407.116.425,00.

A diferença entre o valor total da dívida ativa (R\$3.757.540.294,45) e o valor efetivamente liquidado das CDA inscritas (R\$2.407.116.425,00) é de R\$1.350.423.869,45. Essa variação pode ser atribuída a alguns fatores. Primeiramente, existem casos de parcelamentos interrompidos, nos quais o contribuinte começou a efetuar os pagamentos, mas acabou se tornando inadimplente antes de liquidar completamente a CDA. Isso significa que, apesar de ter havido recuperação de valores nesses casos, as CDA correspondentes não foram totalmente liquidadas.

Outro fator refere-se aos parcelamentos em aberto, nos quais os contribuintes estão realizando os pagamentos regularmente, mas ainda não quitaram completamente as CDA. Nesses casos, o Estado também conseguiu recuperar valores, mas ainda não ocorreu a liquidação das respectivas CDA.

Os valores apresentados na linha "demais recuperados" representam os montantes recuperados referentes a impostos, taxas e multas diferentes do ICMS. Essas cifras englobam os valores recuperados para imposto sobre a propriedade de veículos automotores (IPVA), imposto sobre transmissão causa mortis e doação (ITD), imposto sobre transmissão de bens imóveis (ITBI), bem como taxas e multas diversas.

Tabela 7: Composição dos valores recuperados e liquidados da dívida ativa total ao longo dos anos (2018-2022)

Categoria	2022	2021	2020	2019	2018
QTD CDA	10.535	7.846	2.583	4.944	4.191
ICMS					
liquidado	190.099.708,82	1.106.529.245,92	116.731.155,47	139.543.394,23	854.212.746,27
ICMS recuperado	606.154.664,73	1.394.591.596,58	313.010.228,49	409.473.903,95	1.034.309.920,70
Demais recuperados	151.859.197,43	145.322.494,25	104.807.520,81	158.845.031,25	115.146.930,38
Total recuperado	757.743.862,16	1.539.914.070,83	417.817.749,30	568.318.935,20	1.149.456.851,08

Fonte: O autor, 2023.

A Tabela 8 e a Figura 4 apresentam a participação na recuperação e liquidação da dívida ativa por categoria ao longo dos anos de 2018 a 2022. Os dados demonstram a porcentagem de recuperação em relação ao valor total para duas categorias distintas: ICMS e demais recuperados.

No que diz respeito ao ICMS, observa-se que a participação na recuperação do valor total foi de 79,99% em 2022, 90,56% em 2021, 74,92% em 2020, 72,05% em 2019 e 89,98% em 2018. Isso indica que o ICMS foi responsável por uma parcela significativa da recuperação da dívida ativa ao longo desses anos.

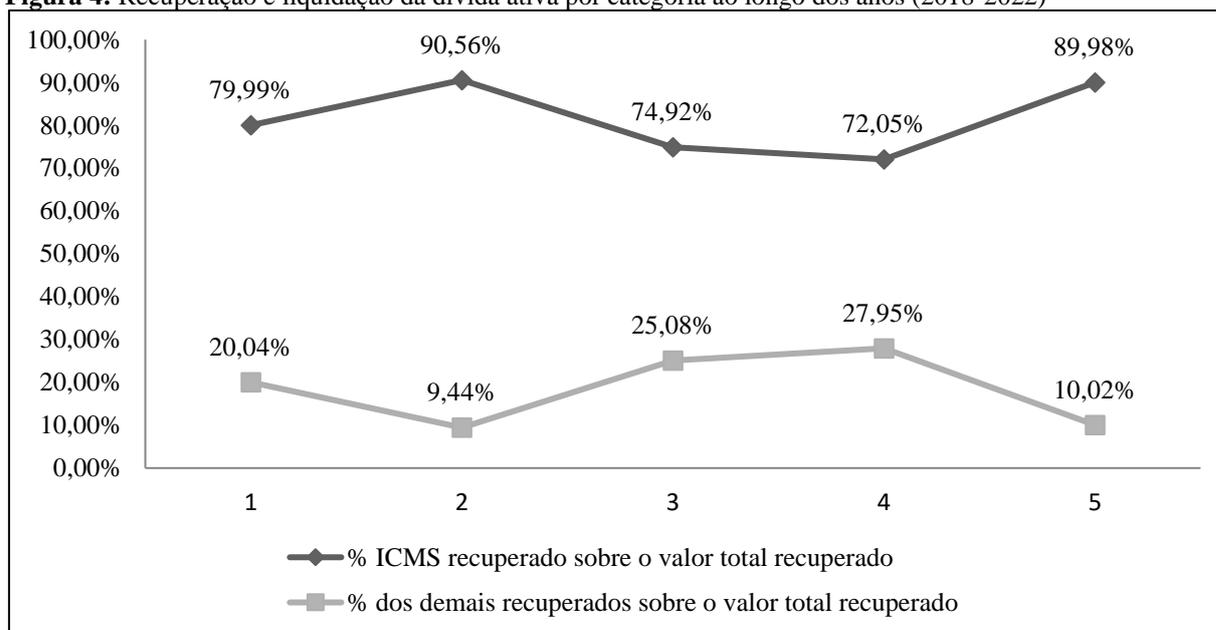
Por outro lado, a categoria dos demais recuperados, que engloba impostos, taxas e multas distintos do ICMS, teve uma participação menor, mas ainda relevante. Em 2022, representou 20,04% da recuperação total, enquanto em 2021 foi de 9,44%, em 2020 foi de 25,08%, em 2019 foi de 27,95% e em 2018 foi de 10,02%.

Tabela 8: Participação na recuperação e liquidação da dívida ativa por categoria ao longo dos anos (2018-2022)

Categoria	2022	2021	2020	2019	2018
% ICMS recuperado sobre o valor total recuperado	79,99%	90,56%	74,92%	72,05%	89,98%
% dos demais recuperados sobre o valor total recuperado	20,04%	9,44%	25,08%	27,95%	10,02%

Fonte: O autor, 2023.

Figura 4: Recuperação e liquidação da dívida ativa por categoria ao longo dos anos (2018-2022)



Fonte: O autor, 2023.

A Tabela 9 e a Figura 5 retratam a variação do total da dívida Ativa ao longo dos anos, bem como a proporção do valor recuperado em relação ao estoque. Em 2022, o total do estoque da dívida ativa foi de R\$135.793.207.387,76, dos quais R\$757.743.862,16 foram recuperados, representando 0,56% do total. No ano anterior, em 2021, o total do estoque foi de R\$121.977.609.267,40, com um total recuperado de R\$1.539.914.070,83, o que equivale a 1,26% do total.

No ano de 2020, o total do estoque foi de R\$118.439.186.473,93, com um valor recuperado de R\$417.817.749,30, correspondendo a 0,35% do total. Em 2019, o estoque da dívida ativa totalizou R\$114.296.760.757,20, sendo recuperados R\$568.318.935,20, que representa 0,50% do total. Em 2018, o estoque da dívida ativa atingiu R\$106.750.103.335,50, com um valor recuperado de R\$1.149.456.851,08, o equivalente a 1,08% do total.

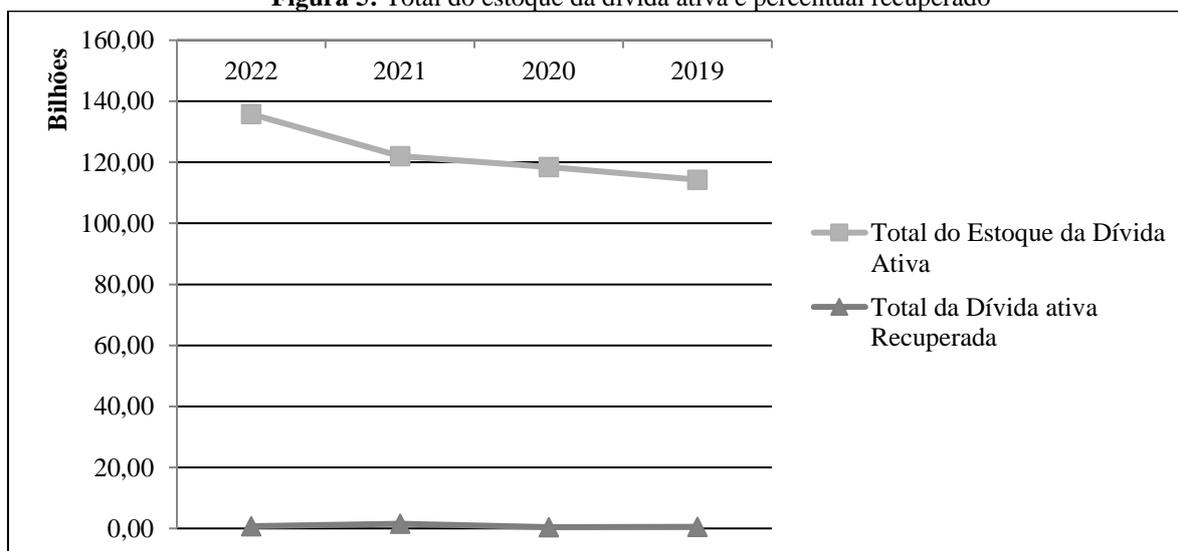
Uma análise longitudinal sobre percentual do estoque da dívida ativa recuperada em relação ao total revela que em 2022 teve uma redução de 0,70% em relação a 2021. Em 2020, houve uma diminuição de 0,91% em relação a 2019, enquanto em 2018, ocorreu um aumento de 0,48% em relação a 2019.

Tabela 9: Total do estoque da dívida ativa e percentual recuperado

Ano	Total do Estoque da dívida ativa	Total do Estoque da dívida ativa recuperada	% do Estoque da dívida ativa recuperada
2022	135.793.207.387,76	757.743.862,16	0.56%
2021	121.977.609.267,40	1.539.914.070,83	1.26%
2020	118.439.186.473,93	417.817.749,30	0.35%
2019	114.296.760.757,20	568.318.935,20	0.50%
2018	106.750.103.335,50	1.149.456.851,08	1.08%

Fonte: O autor, 2023.

Figura 5: Total do estoque da dívida ativa e percentual recuperado



Fonte: O autor, 2023.

Portanto, Através da análise dos dados da dívida ativa do Estado do Rio de Janeiro entre 2018 e 2022, foi possível levantar o montante dos créditos recuperados ao longo desse período, bem como identificar a composição do estoque da dívida ativa e o crescimento do valor do

ICMS, que foi o principal componente da dívida ativa. Além disso, foi possível verificar a proporção de recuperação em relação ao estoque total da dívida ativa ao longo dos anos. Os resultados apresentados atendem ao **objetivo de pesquisa 1, que é "levantar o montante dos créditos recuperados e não recuperados nos últimos 5 anos."**

3.2 Resultados do classificador XGBoost

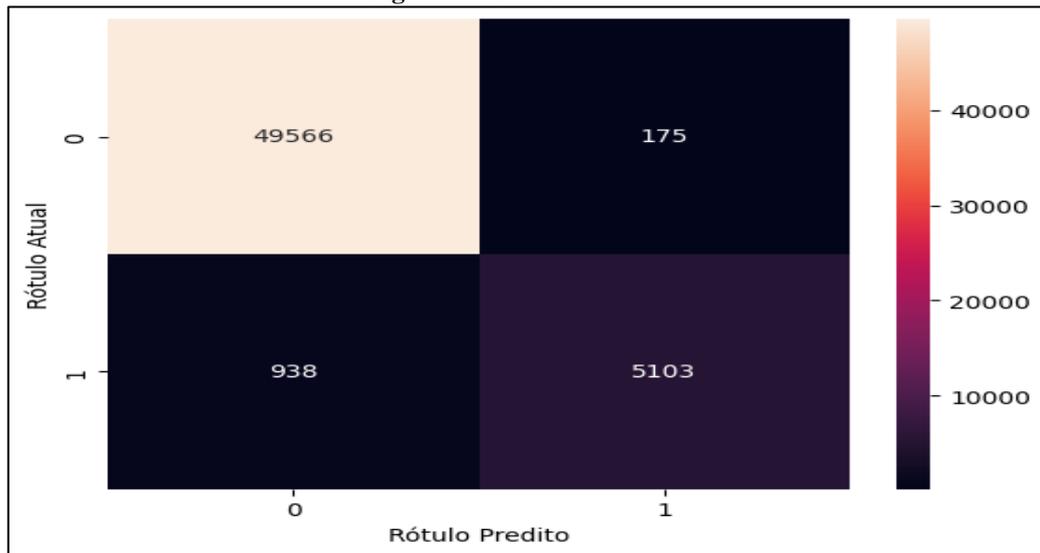
Inicialmente, a base de dados foi segmentada nos conjuntos: amostra de treinamento e amostra de teste. A segmentação da base de dados tem como objetivo permitir a avaliação da capacidade de generalização do classificador diante de dados não observados durante o treinamento.

No presente caso, foram utilizados 223.125 exemplos no conjunto de treinamento. Essa porção dos dados foi empregada para treinar o modelo, permitindo que ele ajustasse seus parâmetros e aprendesse os padrões presentes nos dados das CDA e variáveis explicativas, o total de exemplos no conjunto de treinamento representa 80% da base de dados.

Por outro lado, o conjunto de teste foi composto por 55.782 exemplos. Esses dados não foram utilizados durante o processo de treinamento e foram reservados para avaliar o desempenho final do modelo em dados não observados anteriormente e esse modelo de teste representa 20% da base de dados.

Após o treinamento, o XGBoost foi aplicado às 55.782 CDA do conjunto de teste para avaliação das seguintes métricas: matriz de confusão, precisão (*precision*), revocação (*recall*), *f-1 score*, acurácia (*accuracy*), curva ROC e AUC.

A matriz de confusão resultante é apresentada na Figura 6, na qual são apresentados os totais de acertos e erros de classificação das CDA do conjunto de teste: TP = 5.103, TN = 49.566, FP = 175 e FN = 938. Note que entre as 6.041 CDA recuperadas, 5.103 (84%) foram classificadas corretamente. Já entre as 49.641 CDA não recuperadas 49.566 (99,8%) foram classificadas corretamente. A precisão e o recall alcançados foram da ordem de 97% e 84% respectivamente, resultando em F1 score de 90%. Ademais, a classificação do XGBoost alcançou uma acurácia de 98%.

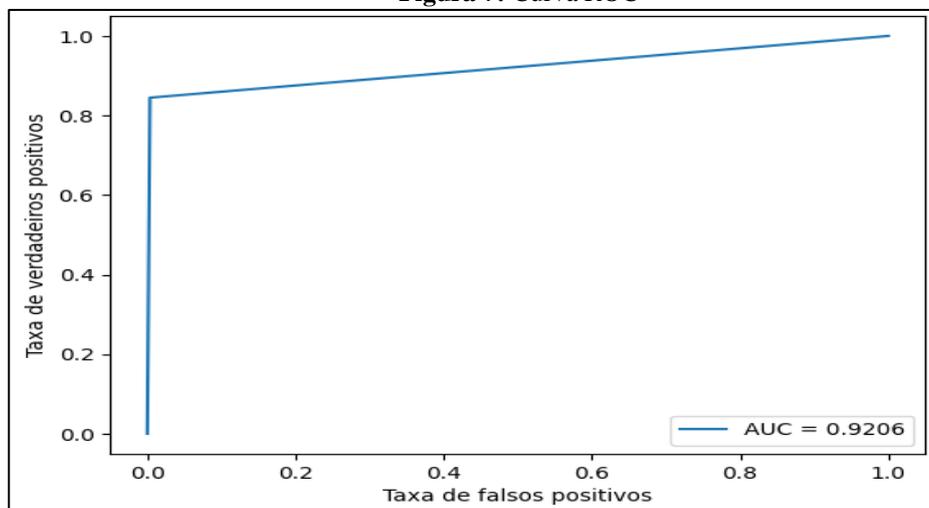
Figura 6: Matriz de confusão

Fonte: O autor, 2023.

Portanto, os resultados mostram que o XGBoost apresentou um desempenho geral satisfatório.

O excelente desempenho alcançado mostra que o classificador XGBoost é confiável em suas predições e tem a capacidade de realizar distinções precisas entre CDA recuperadas e não recuperadas.

Outro resultado obtido na avaliação realizada é a curva ROC ilustrada na Figura 7, cuja Área sob a Curva (AUC) foi de 0,9206, indicando o bom desempenho do XGBoost, i.e., alta sensibilidade e baixa taxa de falsos positivos.

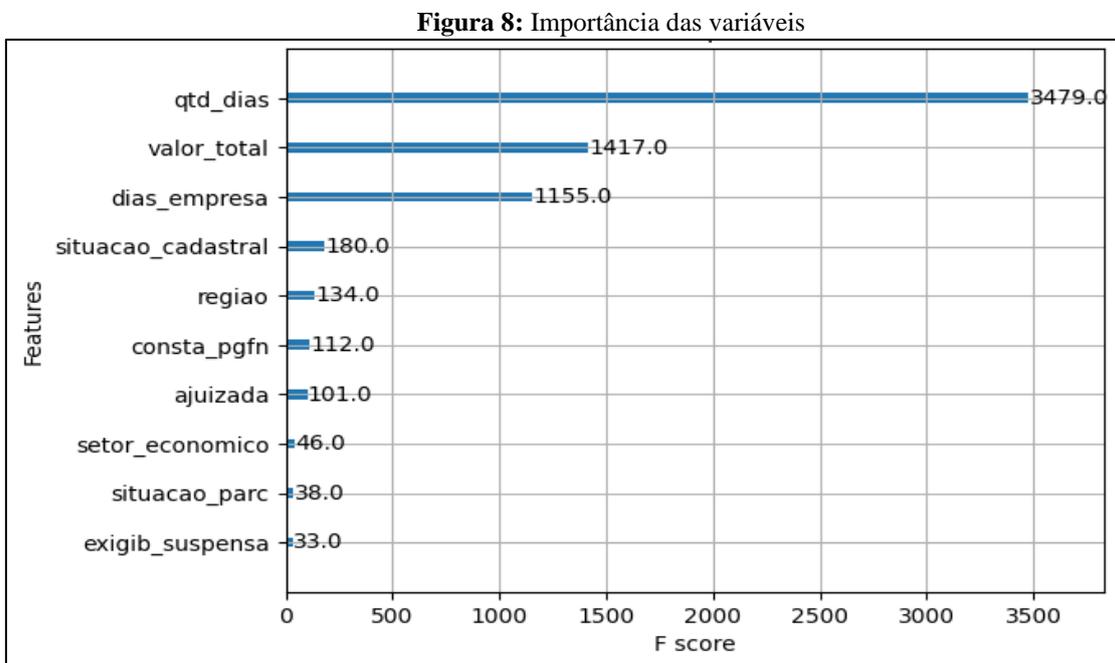
Figura 7: Curva ROC

Fonte: O autor, 2023.

Este resultado reforça a confiabilidade e eficácia do XGBoost classificação das CDA.

O grau de importância de cada variável explicativa (*features*) para a classificação das CDA é ilustrada pelo gráfico de barras na Figura 8. O gráfico permite identificar quais características com maior poder em discriminar as CDA não recuperadas e recuperadas. Um resultado que pode ajudar a ganhar *insights* sobre os fatores que mais influenciam a recuperação de uma CDA.

Conforme indicado na Figura 8, a variável quantidade de dias é a mais importante, seguida respectivamente das seguintes variáveis: valor total, dias de empresa, situação cadastral, região, consta na PGFN, ajuizada, setor econômico, situação do parcelamento e exigibilidade suspensa.



Fonte: O autor, 2023.

Cabe destacar que o *F score* apresentado na Figura 8, mede quantas vezes uma variável é escolhida para fazer divisões da árvore de decisão durante o processo de construção do modelo. Variáveis que são frequentemente escolhidas para divisões são consideradas mais importantes (XIANGJUN et al., 2019).

Com essas informações, pode-se concluir que é importante concentrar esforços e atenção especial na variável "quantidade de dias", a fim de garantir que ela não ultrapasse a média de dias de recuperação, assim como nas variáveis "valor total" e "dias de empresa" que também desempenham um papel significativo na previsão da recuperação de CDA. Isso pode ajudar a melhorar o desempenho do modelo e aumentar a taxa de recuperação de CDA.

É possível que a metodologia atual priorize outras variáveis, como a variável ajuizada e a exigibilidade suspensa, por considerá-las mais determinantes ou por estarem mais alinhadas com as políticas adotadas pela instituição. No entanto, é importante que a seleção das variáveis seja baseada em critérios consistentes e em análises empíricas para garantir a eficácia do modelo de análise da Dívida Ativa.

No entanto, o modelo gerado por este estudo descreve variáveis mais importantes, como os dias de existência da empresa e a situação cadastral das empresas, que talvez tenham grau de relevância devido no modelo atual de análise da Dívida Ativa. As variáveis mencionadas podem, por exemplo, indicar a estabilidade financeira e a regularidade do devedor, assim como a quantidade de dias de inscrição em dívida ativa e o valor total médio da dívida, podem ser relevantes para a análise, pois fornecem informações sobre a história e o valor das dívidas em questão.

Portanto, a metodologia atual pode ter um maior grau de subjetividade na análise dos créditos da Dívida Ativa visto que a definição da dificuldade e da faixa de percentual muitas vezes depende da interpretação do gestor responsável pela análise. A subjetividade não necessariamente é algo negativo, pois permite a flexibilidade na análise e no tratamento dos créditos. No entanto, é essencial que haja critérios claros e consistentes para garantir uma análise coerente dos créditos da Dívida Ativa.

A aplicação do XGBoost treinado ao conjunto de 278.907 CDA da base (248.808 não recuperadas e 30.099 recuperadas) seguida pela classificação das CDA nas categorias A, B, C, D,E,F e G do *rating* do potencial de credibilidade, resultou na Tabela 10 a seguir.

Tabela 10: Faixa de Classificação (*Rating*)

Nível	Faixa de classificação (<i>Rating</i>)	Quantidade de CDA não recuperadas	Quantidade de CDA recuperadas
A	85,00% A 100%	96	22.363
B	70,00% a 84,99%	141	2.134
C	55,00% a 69,99%	263	1.229
D	40,00% a 54,99%	633	997
E	25,00% a 39,99%	1.634	998
F	10,00% a 24,99%	7.471	1.213
G	0,01% a 9,99%	238.570	1.165
	Total	248.808	30.099

Fonte: O autor, 2023.

A Tabela 10 revela a distribuição dos créditos de acordo com a probabilidade de recuperação e, conforme esperado, a maior parte das CDA recuperadas concentram-se nas

categorias A e B, enquanto a maioria das CDA não recuperadas encontram-se nas categorias F e G.

Também vale analisar a distribuição das variáveis explicativas ao longo das categorias do *rating* de potencial de recuperação das CDA. Por exemplo, a Tabela 11 apresenta para cada categoria do *rating* a respectiva média da quantidade de dias de inscrição das CDA, considerando a data base final de inscrição em 31/12/2022. A informação revela que a variável de tempo de inscrição influencia na probabilidade de recuperação, pois quanto maior o tempo de inscrição, menor é a probabilidade de recuperação da dívida.

Tabela 11: Classificação das CDA por faixa de *rating* e média de dias de recuperabilidade

Faixa de classificação (<i>Rating</i>)	Média de dias das CDA não recuperadas	Média de dias das CDA recuperadas
A	1.562	721
B	1.434	1.431
C	1.653	1.933
D	1.971	2.284
E	2.213	2.561
F	2.316	2.915
G	3.060	3.491
Total	14.208	15.336

Fonte: O autor, 2023.

Portanto, é possível notar que as faixas de classificação com menor quantidade de CDA (A, B e C) possuem menor quantidade de dias e menor média de dias de inscrição, indicando um menor tempo de inadimplência. Por outro lado, as faixas de classificação com maior quantidade de CDA (E, F e G) apresentam uma maior quantidade de dias e média de dias, indicando um maior tempo de inadimplência. A categoria G, em particular, possui a maior quantidade de CDA, e a maior média de dias, o que pode representar uma situação de inadimplência mais crítica.

Fazendo a mesma análise para a variável “Valor Total”, i.e., o valor total de cada uma das CDA inscritas em dívida ativa, considerando a data base de 31/12/2022, obteve-se a Tabela 12 com o valor médio da dívida em cada categoria do *rating*.

O resultado aponta que quanto maior o valor médio das dívidas, pior é a probabilidade de recuperação.

Tabela 12: Classificação das CDA por *rating*, quantidade de CDA, valor total e valor médio

Faixa de classificação (<i>Rating</i>)	Quantidade das CDA	Valor Total (em R\$)	Valor médio (em R\$)
A	96	1.078.617,00	11.236,00
B	141	9.310.857,00	66.034,00
C	263	5.523.514,00	21.002,00
D	633	65.856.387,00	104.039,00
E	1.634	348.346.602,00	213.186,00
F	7.471	2.299.114.065,00	307.738,00
G	238.570	127.245.581.933,00	533.368,00
Total	248.808	129.974.811.975,00	1.256.603,00

Fonte: O autor, 2023.

No entanto, cabe registrar uma exceção a esse padrão de recuperabilidade: a faixa de classificação "B" possui um valor médio das dívidas maior do que a faixa "C". Essa discrepância pode ser atribuída a fatores específicos, como a presença de dívidas individualmente significativas ou uma combinação de diferentes fatores.

Em qualquer caso, a análise dos valores médios das dívidas por faixa de *rating* fornece informações sobre a probabilidade de recuperação, que podem ser usadas para orientar estratégias de recuperação de dívidas.

A variável “dias de empresa” apresenta o tempo de existência da empresa, considerando a data de sua abertura informada no cartão do CPNJ, até a data base de 31/12/2022. A Tabela 13 apresenta a classificação das CDA por *rating* e tempo médio de existência das empresas em dias.

Tabela 13: Classificação das CDA por *rating* e tempo médio de existência das empresas em dias

Faixa de classificação (<i>Rating</i>)	Quantidade de CDA	Dias de existência das empresas	Tempo médio de existência das empresas (em dias)
A	96	711.053	7.407
B	141	1.139.189	8.079
C	263	1.940.628	7.379
D	633	4.942.076	7.807
E	1.634	13.409.448	8.207
F	7.471	61.758.349	8.266
G	238.570	1.853.617.194	7.770
Total	248.808	1.937.517.937	7.787

Fonte: O autor, 2023.

Apesar dos valores nos níveis B, C e G na Tabela 13, os resultados apontam para uma tendência de menor probabilidade de recuperação com o avanço da idade da empresa.

Conforme indicado na Tabela 14, a região metropolitana possui a maior participação em todos os níveis de rating.

Tabela 14: Classificação das CDA por região em valores absolutos

Faixa de classificação (<i>Rating</i>)	Região metropolitana		Interior do Rio de Janeiro		Outros Estados		Total
	n(=)	%	n(=)	%	n(=)	%	
A	60	63,00%	26	27,00%	10	10,00%	96
B	79	56,00%	40	28,00%	22	16,00%	141
C	171	65,00%	54	21,00%	38	14,00%	263
D	422	67,00%	131	21,00%	80	13,00%	633
E	1129	69,00%	323	20,00%	182	11,00%	1634
F	4883	65,00%	1411	19,00%	1177	16,00%	7471
G	162345	68,00%	49016	21,00%	27209	11,00%	238570

Fonte: O autor, 2023.

Já a Tabela 15 mostra que, quanto melhor o *rating* de recuperação, menor é o percentual de ajuizamento, o que sugere que, outras formas de resolução da cobrança, que não o processo judicial, possa ter melhor probabilidade de recuperação dos créditos. Destaca-se ainda o percentual de CDA no nível G do *rating* que chega a 50%.

Tabela 15: Classificação das CDA por débitos ajuizados

Faixa de classificação (<i>Rating</i>)	Quantidade de CDA	Ajuizada	%
A	96	27	28,00%
B	141	34	24,00%
C	263	74	28,00%
D	633	200	32,00%
E	1634	635	39,00%
F	7471	3.284	44,00%
G	238570	118.618	50,00%
Total	248.808	122.872	

Fonte: O autor, 2023.

Portanto, a presente subseção atende ao **objetivo de pesquisa 3 que consiste em “elaborar modelo de classificação dos créditos (rating) que permita o direcionamento dos esforços da PGE/RJ na busca de créditos com maior probabilidade de recuperação”**.

3.2.1 Faixa de classificação (*Rating*) A

O nível A refere-se a faixa de classificação mais alta que vai de 85,00% a 100,00%, logo, apresenta os créditos com melhor probabilidade de recuperação e totaliza o montante de R\$ 1.078.617,00. O resultado das variáveis estudadas nessa faixa constam na Tabela 16.

Tabela 16: Variáveis das Certidões de Dívida Ativa (CDA) - *rating* A (85,00% a 100,00%)

Variável analisada com relação ao <i>rating</i> A	Valor
Quantidade CDA	96
V ₂ - Dívida ajuizada	28,00%
V ₃ - Exigibilidade suspensa	2,00%
V ₉ – Existência de dívida na PGFN	23,00%
V ₈ - Média de dias de existência da empresa	7.407 dias
V ₈ - Média da quantidade de dias (tempo inscrito em Dívida Ativa)	1.562 dias
V ₁ - Valor total médio das dívidas inscritas	R\$ 11.236,00
V _{criada} – Média da probabilidade de recebimento do <i>rating</i> A	92,00%
V ₇ – Quantidade de CNPJ com situação cadastral nula	0
V ₇ - Quantidade de CNPJ com situação cadastral ativa	68
V ₇ - Quantidade de CNPJ com situação cadastral suspensa	0
V ₇ - Quantidade de CNPJ com situação cadastral inapta	5
V ₇ - Quantidade de CNPJ com situação cadastral baixada	16
V ₇ - Quantidade de CNPJ com situação cadastral não encontrada	7
V ₆ - Região metropolitana	63,00%
V ₆ - Interior do RJ	27,00%
V ₆ - Outros estados	10,00%
V ₁₀ - Setor econômico primário	3,00%
V ₁₀ - Setor econômico secundário	15,00%
V ₁₀ - Setor econômico terciário	75,00%
V ₁₀ - Setor econômico não encontrado	7,00%
V ₄ - Com Parcelamento	0,00%
V ₄ - Sem Parcelamento	100,00%

Fonte: O autor, 2023.

Essa faixa engloba os devedores que apresentam as melhores condições de quitação dos créditos da dívida ativa dentro do grupo. A média da quantidade de dias em que a dívida está inscrita é de 1.562 dias, enquanto a média de dias para recuperação das CDA nesse nível de *rating* foi de 721 dias (tabela 11), o que demonstra a necessidade de atuação da PGE/RJ, nessas CDA com menor tempo de inscrição em dívida ativa.

A média de dias de existência da empresa é de 7.407 dias e o valor total médio das CDA é de R\$ 11.236,00. Constata-se também que a probabilidade média de recuperação das CDA nesse nível é de 92%.

3.2.2 Faixa de classificação (*Rating*) B

O nível B, abrange a faixa com percentuais de recuperação que vai de 70,00% a 84,99%, é o segundo melhor nível do modelo em relação à probabilidade de recuperação. Observa-se uma proporção menor em relação ao nível anterior, e totaliza o montante de R\$ 9.310.857,00. O resultado das variáveis estudadas nessa faixa consta na Tabela 17.

Tabela 17: Variáveis das Certidões de Dívida Ativa (CDA) referente ao *rating* B (70,00% a 84,99%)

Variável analisada com relação ao <i>rating</i> B	Valor
Quantidade CDA	141
V ₂ - Dívida ajuizada	24,00%
V ₃ - Exigibilidade suspensa	2,00%
V ₉ - Existência de dívida na PGFN	41,00%
V ₈ - Média de dias de existência da empresa	8.079 dias
V ₈ - Média da quantidade de dias (tempo inscrito em Dívida Ativa)	1.434 dias
V ₁ - Valor total médio das dívidas inscritas	R\$ 66.034,00
Variada - Média da probabilidade de recebimento do <i>rating</i> B	76%
V ₇ - Quantidade de CNPJ com situação cadastral nula	0
V ₇ - Quantidade de CNPJ com situação cadastral ativa	109
V ₇ - Quantidade de CNPJ com situação cadastral suspensa	0
V ₇ - Quantidade de CNPJ com situação cadastral inapta	5
V ₇ - Quantidade de CNPJ com situação cadastral baixada	22
V ₇ - Quantidade de CNPJ com situação cadastral não encontrada	5
V ₆ - Região metropolitana	56,00%
V ₆ - Interior do RJ	28,00%
V ₆ - Outros estados	16,00%
V ₁₀ - Setor econômico primário	1,00%
V ₁₀ - Setor econômico secundário	17,00%
V ₁₀ - Setor econômico terciário	78,00%
V ₁₀ - Setor econômico não encontrado	4,00%
V ₄ - Com Parcelamento	0,00%
V ₄ - Sem Parcelamento	100,00%

Fonte: O autor, 2023.

Os resultados apresentados demonstram as dívidas com a segunda melhor expectativa de recebimentos dentre os níveis estudados. Destacam-se, a quantidade de CDA inscrita que é mais que o dobro do nível anterior e o valor total médio da dívida que é de R\$66.034,30.

3.2.3 Faixa de classificação (*Rating*) C

O nível C, abrange a faixa com percentuais de recuperação que vai de 55,00% a 69,99%, é o terceiro melhor nível do modelo em relação à probabilidade de recuperação. Observa-se que o valor total desse nível é menor que o anterior e totaliza o montante de R\$ 5.523.514,00. O resultado das variáveis estudadas nessa faixa consta na Tabela 18.

Tabela 18: Variáveis das Certidões de Dívida Ativa (CDA) referente ao *rating* C (55,00% a 69,99%)

Variável analisada com relação ao <i>rating</i> C	Valor
Quantidade CDA	263
V ₂ - Dívida ajuizada	28,00%
V ₃ - Exigibilidade suspensa	5,00%
V ₉ - Existência de dívida na PGFN	31,00%
V ₈ - Média de dias de existência da empresa	7.379 dias
V ₈ - Média da quantidade de dias (tempo inscrito em Dívida Ativa)	1.653 dias
V ₁ - Valor total médio das dívidas inscritas	R\$ 21.002,00
Variada - Média da probabilidade de recebimento do <i>rating</i> C	61,00%
V ₇ - Quantidade de CNPJ com situação cadastral nula	0
V ₇ - Quantidade de CNPJ com situação cadastral ativa	201
V ₇ - Quantidade de CNPJ com situação cadastral suspensa	4
V ₇ - Quantidade de CNPJ com situação cadastral inapta	8
V ₇ - Quantidade de CNPJ com situação cadastral baixada	41
V ₇ - Quantidade de CNPJ com situação cadastral não encontrada	9
V ₆ - Região metropolitana	65,00%
V ₆ - Interior do RJ	21,00%
V ₆ - Outros estados	14,00%
V ₁₀ - Setor econômico primário	0,00%
V ₁₀ - Setor econômico secundário	12,00%
V ₁₀ - Setor econômico terciário	84,00%
V ₁₀ - Setor econômico não encontrado	3,00%
V ₄ - Com Parcelamento	0,00%
V ₄ - Sem Parcelamento	100,00%

Fonte: O autor, 2023.

Esse resultado sugere que os devedores nessa faixa possuem uma classificação mais moderada em comparação com os níveis superiores (A e B), indicando uma probabilidade moderada de recuperação dos créditos. Destaca-se a alta participação do setor terciário que chega a 84%.

3.2.4 Faixa de classificação (*Rating*) D

O nível D compreende a faixa com percentuais de recuperação de probabilidade que vão de 40,00% a 54,99%, observa-se um valor maior total inscrito que é quase 12 vezes maior que o nível anterior e totaliza o montante de R\$ 65.856.387,00. O resultado das variáveis estudadas nessa faixa consta na Tabela 19.

Tabela 19: Variáveis das Certidões de Dívida Ativa (CDA) referente ao *rating* D (40,00% a 54,99%)

Variável analisada com relação ao <i>rating</i> D	Valor
Quantidade CDA	633
V ₂ - Dívida ajuizada	32,00%
V ₃ - Exigibilidade suspensa	5,00%
V ₉ - Existência de dívida na PGFN	29,00%
V ₈ - Média de dias de existência da empresa	7.807 dias
V ₈ - Média da quantidade de dias (tempo inscrito em Dívida Ativa)	1.971 dias
V ₁ - Valor total médio das dívidas inscritas	R\$ 104.039,00
Variada - Média da probabilidade de recebimento do <i>rating</i> D	46,00%
V ₇ - Quantidade de CNPJ com situação cadastral nula	0
V ₇ - Quantidade de CNPJ com situação cadastral ativa	465
V ₇ - Quantidade de CNPJ com situação cadastral suspensa	6
V ₇ - Quantidade de CNPJ com situação cadastral inapta	19
V ₇ - Quantidade de CNPJ com situação cadastral baixada	124
V ₇ - Quantidade de CNPJ com situação cadastral não encontrada	19
V ₆ - Região metropolitana	67,00%
V ₆ - Interior do RJ	21,00%
V ₆ - Outros estados	13,00%
V ₁₀ - Setor econômico primário	1,00%
V ₁₀ - Setor econômico secundário	13,00%
V ₁₀ - Setor econômico terciário	83,00%
V ₁₀ - Setor econômico não encontrado	3,00%
V ₄ - Com Parcelamento	0,00%
V ₄ - Sem Parcelamento	100,00%

Fonte: O autor, 2023.

Esse resultado sugere que os devedores nessa faixa possuem uma classificação mais difícil em comparação com os níveis superiores (A, B e C), indicando uma probabilidade média de recuperação dos créditos inscritos. Destaca- que o valor médio devido por CDA ultrapassa pela primeira vez o valor de R\$100.000,00 (cem mil reais), considerando os níveis de *rating* A, B e C.

3.2.5 Faixa de classificação (*Rating*) E

O nível D compreende a faixa que abrange de 25,00% a 39,99% dos percentuais de recuperação de probabilidade e totaliza o montante de R\$ 65.856.387,00. O resultado das variáveis estudadas nessa faixa consta na Tabela 20.

Tabela 20: Variáveis das Certidões de Dívida Ativa (CDA) referente ao *rating* E (25,00% a 39,99)

Variável analisada com relação ao <i>rating</i> E	Valor
Quantidade CDA	1634
V ₂ - Dívida ajuizada	39,00%
V ₃ - Exigibilidade suspensa	5,00%
V ₉ - Existência de dívida na PGFN	29,00%
V ₈ - Média de dias de existência da empresa	8.207 dias
V ₈ - Média da quantidade de dias (tempo inscrito em Dívida Ativa)	2.213 dias
V ₁ - Valor total médio das dívidas inscritas	R\$ 213.186,00
Variada - Média da probabilidade de recebimento do <i>rating</i> E	31,00%
V ₇ - Quantidade de CNPJ com situação cadastral nula	0
V ₇ - Quantidade de CNPJ com situação cadastral ativa	1136
V ₇ - Quantidade de CNPJ com situação cadastral suspensa	14
V ₇ - Quantidade de CNPJ com situação cadastral inapta	92
V ₇ - Quantidade de CNPJ com situação cadastral baixada	326
V ₇ - Quantidade de CNPJ com situação cadastral não encontrada	66
V ₆ - Região metropolitana	69,00%
V ₆ - Interior do RJ	20,00%
V ₆ - Outros estados	11,00%
V ₁₀ - Setor econômico primário	1,00%
V ₁₀ - Setor econômico secundário	14,00%
V ₁₀ - Setor econômico terciário	80,00%
V ₁₀ - Setor econômico não encontrado	4,00%
V ₄ - Com Parcelamento	0,00%
V ₄ - Sem Parcelamento	100,00%

Fonte: O autor, 2023.

Nessa faixa, os devedores possuem uma classificação mais baixa em relação aos níveis anteriores (A, B, C e D). Esse resultado sugere uma probabilidade baixa de recuperação dos créditos inscritos. Observa-se que a dívida média por CDA de R\$213.186,00 é mais que o dobro do nível anterior.

3.2.6 Faixa de classificação (Rating) F

O nível F compreende a faixa que abrange de 10,00% a 24,99% dos percentuais de recuperação de probabilidade e totaliza o montante de R\$ 2.299.114.065,00. O resultado das variáveis estudadas nessa faixa consta na Tabela 21.

Tabela 21: Variáveis das Certidões de Dívida Ativa (CDA) referente ao *rating* F (10,00% a 24,99%)

Variável analisada com relação ao rating F	Valor
Quantidade CDA	7471
V ₂ - Dívida ajuizada	44%
V ₃ - Exigibilidade suspensa	4%
V ₉ - Existência de dívida na PGFN	33%
V ₈ - Média de dias de existência da empresa	8.266
V ₈ - Média da quantidade de dias (tempo inscrito em Dívida Ativa)	2.316
V ₁ - Valor total médio das dívidas inscritas	R\$ 307.738,00
V _{criada} - Média da probabilidade de recebimento do rating F	15%
V ₇ - Quantidade de CNPJ com situação cadastral nula	6
V ₇ - Quantidade de CNPJ com situação cadastral ativa	5187
V ₇ - Quantidade de CNPJ com situação cadastral suspensa	59
V ₇ - Quantidade de CNPJ com situação cadastral inapta	438
V ₇ - Quantidade de CNPJ com situação cadastral baixada	1518
V ₇ - Quantidade de CNPJ com situação cadastral não encontrada	263
V ₆ - Região metropolitana	65%
V ₆ - Interior do RJ	19%
V ₆ - Outros estados	16%
V ₁₀ - Setor econômico primário	1%
V ₁₀ - Setor econômico secundário	15%
V ₁₀ - Setor econômico terciário	80%
V ₁₀ - Setor econômico não encontrado	4%
V ₄ - Com Parcelamento	0%
V ₄ - Sem Parcelamento	100%

Fonte: O autor, 2023.

Nessa faixa, os devedores possuem uma classificação mais baixa em relação aos níveis anteriores e esse fato implica em uma probabilidade muito baixa de recuperação dos créditos inscritos.

3.2.7 Faixa de classificação (*Rating*) G

O nível G compreende a faixa que abrange de 0,01% a 9,99% dos percentuais de recuperação de probabilidade e totaliza o montante de R\$127.245.581.933,00, é a pior faixa do rating. O resultado das variáveis estudadas nessa faixa consta na Tabela 22.

Tabela 22: Variáveis das Certidões de Dívida Ativa (CDA) referente ao *rating* G (0,01% a 9,99%)

Variável analisada com relação ao <i>rating</i> G	Valor
Quantidade CDA	238570
V ₂ - Dívida ajuizada	50,00%
V ₃ - Exigibilidade suspensa	1,00%
V ₉ - Existência de dívida na PGFN	40,00%
V ₈ - Média de dias de existência da empresa	7.770 dias
V ₈ - Média da quantidade de dias (tempo inscrito em Dívida Ativa)	3.060 dias
V ₁ - Valor total médio das dívidas inscritas	R\$ 533.368,00
Variada - Média da probabilidade de recebimento do <i>rating</i> G	1,00 %
V ₇ - Quantidade de CNPJ com situação cadastral nula	64
V ₇ - Quantidade de CNPJ com situação cadastral ativa	91591
V ₇ - Quantidade de CNPJ com situação cadastral suspensa	2694
V ₇ - Quantidade de CNPJ com situação cadastral inapta	75000
V ₇ - Quantidade de CNPJ com situação cadastral baixada	56228
V ₇ - Quantidade de CNPJ com situação cadastral não encontrada	12993
V ₆ - Região metropolitana	68,00%
V ₆ - Interior do RJ	21,00%
V ₆ - Outros estados	11,00%
V ₁₀ - Setor econômico primário	1,00%
V ₁₀ - Setor econômico secundário	16,00%
V ₁₀ - Setor econômico terciário	78,00%
V ₁₀ - Setor econômico não encontrado	6,00%
V ₄ - Com Parcelamento	4,00%
V ₄ - Sem Parcelamento	96,00%

Fonte: O autor, 2023.

Nessa faixa, os devedores possuem uma classificação mais baixa em relação aos níveis existentes (A, B, C, D, E e F). Isso implica no fato de que as CDA inscritas nessa faixa possuem as piores probabilidades de recuperação dos créditos inscritos. Ademais, o *rating* G das Certidões de Dívida Ativa (CDA) apresenta uma quantidade significativa de 238.570 CDA nessa faixa de classificação.

Destaca-se que 91.151 CDA apresentam situação cadastral ativa, o que indica que a PGE/RJ deve se concentrar nessas empresas cujo CNPJ indica atividade da empresa. Temos

ainda que metade das CDA deste grupo se encontram ajuizadas indicando maior dificuldade na recuperação.

3.3 Proposta de modelo de ajuste de perdas para Dívida Ativa RJ

Conforme já abordado na subseção 2.3 deste estudo, o MCASP determina que o responsável pela elaboração do ajuste de perdas seja o órgão competente que faça a gestão da Dívida Ativa, neste caso, a Procuradoria Geral do Estado do Rio de Janeiro, todavia, ele não especifica uma metodologia para o cálculo do ajuste de perdas. O manual menciona ainda que a mensuração do ajuste de perda se baseie em estudos especializados que balizem e qualifiquem os créditos inscritos, de modo a não superestimar e nem subavaliar o patrimônio real do ente público.

Um dos objetivos deste trabalho é a proposição de um modelo que permita a elaboração do ajuste de perda da Dívida Ativa do Estado do Rio de Janeiro considerando um viés técnico, portanto, levando-se em conta o arcabouço teórico utilizado neste trabalho e a utilização de um algoritmo de aprendizado de máquina para análise da probabilidade de recuperabilidade dos créditos. Para tanto são sugeridos os seguintes passos:

1. Classificar as CDA de acordo com as faixas de classificação (*rating*).
2. Calcular a probabilidade média de recuperação para cada faixa de classificação. Essa probabilidade média indica a chance média de recuperação das CDA em cada faixa.
3. Multiplicar o valor total de cada CDA pela probabilidade média de recuperação correspondente à sua faixa de classificação para determinar expectativa de recuperação para cada CDA.
4. Somar os valores ajustados de todas as CDA para obter o valor total ajustado, que representa a expectativa de recuperação para o conjunto de CDA analisadas.
5. Comparar o valor total ajustado com o valor total das CDA para determinar a diferença entre o valor esperado de recuperação e o valor original das dívidas. Essa diferença é a base para o lançamento de ajuste para perdas.

Para calcular o valor ajustado de cada CDA com base nas probabilidades de recuperação fornecidas pelo modelo XGBoost, foi multiplicado o valor total da dívida pela probabilidade média correspondente à sua faixa de classificação.

A expectativa de recuperação para o conjunto de CDA analisadas é correspondente à soma dos valores ajustados de cada faixa de classificação e equivale a R\$ 1.767.042.236,22. Considerando que o valor do Estoque da Dívida é de R\$ 129.974.811.975,00, o valor da perda estimada corresponde a R\$ 128.207.769.739,09. Os resultados constam na Tabela 25.

Tabela 23: Estimativa de Perdas para CDA por Faixa de Classificação

Faixa	Valor da Dívida ativa	Probabilidade de recuperabilidade	Valor Recuperável	Estimativa de perda
A	R\$ 1.078.617	92,00%	R\$ 992.327,64	R\$ 86.289,36
B	R\$ 9.310.857	76,00%	R\$ 7.076.251,32	R\$ 2.234.605,68
C	R\$ 5.523.514	61,00%	R\$ 3.369.343,54	R\$ 2.154.170,46
D	R\$ 65.856.387	46,00%	R\$ 30.293.938,02	R\$ 35.562.448,98
E	R\$ 348.346.602	31,00%	R\$ 107.987.446,62	R\$ 240.359.155,38
F	R\$ 2.299.114.065	15,00%	R\$ 344.867.109,75	R\$ 1.954.246.955,25
G	R\$ 127.245.581.933	1,00%	R\$ 1.272.455.819,33	R\$ 125.973.126.113,98
Total	R\$ 129.974.811.975,00		R\$ 1.767.042.236,22	R\$ 128.207.769.739,096

Fonte: O autor, 2023.

Portanto, o percentual de perdas estimado em relação ao valor total da dívida ativa é de aproximadamente 98,64%. Embora o percentual de perdas estimado seja elevado, é importante destacar que esse valor reflete, com base em metodologia, uma estimativa confiável das probabilidades de recuperação e calcular as perdas esperadas.

A metodologia atualmente utilizada na dívida ativa está descrita no item 2.6 deste trabalho e para efeitos de comparação entre os valores dessa metodologia e a proposição deste trabalho, temos que, o valor recuperável dos créditos tributários é de R\$45.955.274.649,69, o que representa aproximadamente 35,35% de valor recuperável em relação ao valor total inscrito de R\$130.126.199.181,60.

Na proposição deste trabalho o valor recuperável estimado é de R\$1.767.042.236,22, que representa aproximadamente 1,36% do valor dos créditos inscritos de ICMS, há de se ressaltar que o atual modelo da dívida ativa considera todo o grupo de créditos tributários, enquanto este trabalho, está considerando somente o ICMS. Cabe ainda destacar que o percentual de recuperabilidade dos créditos aqui estimado, está muito mais próximo dos percentuais de recuperação anuais apresentados na tabela 9 deste trabalho na seção 4.1 de análise dos dados.

Essas estimativas são fundamentais para o planejamento financeiro e a tomada de decisões estratégicas. Ao reconhecer e considerar o potencial de perdas, o gestor e demais partes interessadas pode adotar medidas proativas para mitigar os riscos, como a implementação de estratégias de recuperação de crédito, a negociação de acordos com devedores, o estabelecimento de políticas de controle mais rigorosas, ou até mesmo proposições legislativas.

Cabe mencionar que, em determinadas circunstâncias, como crises econômicas, instabilidade política ou outros fatores externos, as perdas podem ser ainda mais significativas.

Portanto, a presente subseção atende ao **objetivo de pesquisa 4 que consiste em “constituir uma classificação da probabilidade da recuperabilidade dos créditos inscritos em dívida ativa, permitindo o ajuste dos valores para perdas ou continuidade como provável recebimento”**.

CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS

O objetivo deste trabalho foi contribuir para o aprimoramento da gestão da dívida ativa do Estado do Rio de Janeiro. Para isso, foi proposta uma abordagem que fornece uma classificação da probabilidade de recuperação dos créditos inscritos, permitindo que a Procuradoria Geral do Estado do Rio de Janeiro (PGE/RJ) direcione seus esforços de recuperação de forma mais eficiente e reconheça contabilmente esses créditos de maneira mais objetiva.

A metodologia adotada neste estudo envolveu a análise quantitativa e o uso do algoritmo de aprendizado de máquina conhecido como XGBoost, cuja utilização e aplicação ocorreu no programa *Python* e exigiu a obtenção de conhecimento dessa ferramenta. Através dessa abordagem, foram identificados os fatores que influenciam a probabilidade de recuperação das Certidões de Dívida Ativa (CDA) inscritas, levando em consideração 10 variáveis de estudo, conforme resultado encontrado.

Após a aplicação do modelo, foi possível elencar quais os fatores, ou variáveis, que mais contribuem para a recuperação do créditos inscritos de ICMS em dívida ativa no estado do Rio de Janeiro, dentre as quais, destacam-se a variável “quantidade de dias”, que denota a importância de uma atuação célere da PGE na cobrança dos créditos inscritos, o “valor total”, que representa a necessidade de cobrança das dívidas de menor valor médio, e por fim, a variável “dias empresa” que ressalta a necessidade de concentrar esforços na busca dos créditos inscritos de empresas que estejam com suas dívidas a menos tempo inscritas em dívida ativa.

A importância de cada variável foi elencada por ordem de relevância e com base nos resultados obtidos na análise dessas variáveis, foi estabelecido um *rating* de créditos, proporcionando uma visão mais clara sobre a probabilidade de recuperação de cada CDA e permitindo a priorização dos esforços de recuperação nos casos com maior probabilidade de sucesso. Além disso, esses resultados podem contribuir para uma melhor estimativa de perdas relacionadas à dívida ativa do Estado do Rio de Janeiro.

Atualmente, a PGE/RJ utiliza um método baseado no julgamento profissional dos Procuradores, sem a utilização de critérios estatísticos objetivos, para estimar as perdas com créditos da dívida ativa. No entanto, com base nos resultados do modelo proposto, estima-se que as perdas cheguem a aproximadamente 98,64% do valor total da dívida ativa, revelando a necessidade de uma revisão dessa metodologia.

Os resultados obtidos por meio do modelo de XGBoost podem contribuir significativamente para a construção de uma estimativa mais precisa das perdas com os créditos

da dívida ativa do Estado do Rio de Janeiro. Além disso, esse enfoque mais objetivo e baseado em dados pode auxiliar no desenvolvimento de estratégias para o direcionamento de esforços para recuperação dos créditos de forma mais eficaz, resultando em um melhor desempenho na gestão da Dívida ativa e no retorno financeiro para os cofres públicos.

Como proposta de estudos futuros, sugere-se a aplicação do modelo em outros impostos não abordados nesse trabalho, tais como, IPVA, ITBI e o ITD, sugere-se ainda a continuidade da pesquisa com a expansão do modelo para incluir outras variáveis relevantes, entre as quais podemos citar, a data de vencimento das certidões, tendo em vista que esse trabalho contemplou a data de inscrição dos créditos, além de avaliar sua aplicação em outras Procuradorias e contextos.

REFERÊNCIAS

AGRESTI, Alan; FINLAY, Bárbara. **Métodos estatísticos para as ciências sociais.** (Métodos de pesquisa). Porto Alegre: Grupo A, 2012. *E-book*.

ALVES, F. J. S. **Adesão do contabilista ao código de ética da sua profissão:** um estudo empírico sobre percepções. 2005. 273 f. Tese (Doutorado em Ciências Contábeis) – Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, 2005.

ASTOLFI, Fabio Sortin; SILVA, José Américo Oliveira. **Securitização de créditos de entes subnacionais.** 2021. Dissertação (Mestrado profissional MPGPP) - Escola de Administração de Empresas de São Paulo, Fundação Getúlio Vargas, 2021.

BAESENS, B., VLASSELAER, V.V., VERBEKE, W. **Fraud analytics using descriptive, predictive and social network techniques:** a guide to data science for fraud detection. Nova Jersey: Wiley, 2015.

BARBIERI, Carlos. **Governança de dados.** Rio de Janeiro: Alta Books, 2020. *E-book*.

BARBOSA JÚNIOR, Ney Lopes. **Economia comportamental e regra de cobrança para recuperação de créditos tributários na Cidade de Fortaleza.** 2020. 78f. Dissertação (Mestrado em Economia do Setor Público) - Faculdade de Economia, Administração, Atuária e Contabilidade - FEAAC, Programa de Economia Profissional - PEP, Universidade Federal do Ceará, Fortaleza, 2020.

BOAS, A. A. V.; PEREIRA, J. R.; TEIXEIRA, M. P. dos R.; ZWICK, E. Administração pública tupiniquim: reflexões a partir da Teoria N e da Teoria P de Guerreiro Ramos. **Cadernos EBAPE.BR**, Rio de Janeiro, v. 10, n. 2, p. 284-301, 2012.

BRASIL. **Código de Tributário Nacional.** Lei nº 5. 172, de 25 de Outubro de 1966. Dispõe sobre o Sistema Tributário Nacional e institui normas gerais de direito tributário aplicáveis à União, Estados e Municípios. Brasília: Congresso Nacional. Disponível em www.planalto.gov.br. Acesso em: 4 mar. 2023.

BRASIL. **Lei Complementar n 101, de 4 de maio de 2000.** Estabelece normas de finanças públicas voltadas para a responsabilidade e na gestão fiscal e dá outras providências. Brasília: Congresso Nacional. Disponível em www.planalto.gov.br. Acesso em: 2 mar. 2023.

BRASIL. **Lei Federal nº 4.320, de 17 de março de 1964.** Estatui Normas Gerais de Direito Financeiro para elaboração e controle dos orçamentos e balanços da União, dos Estados, dos Municípios e do Distrito Federal. Brasília: Congresso Nacional. Disponível em www.planalto.gov.br. Acesso em: 2 mar. 2023.

BRASIL. Secretaria do Tesouro Nacional. **Manual de Contabilidade Aplicada ao Setor Público (MCASP)**., 9a Edição. Brasília: STN, 2021.

BRESSER-PEREIRA, Luiz Carlos. Da administração pública burocrática à gerencial. **Revista do Serviço Público**, Brasília, v. 73, 2022. Disponível em: <https://revista.enap.gov.br/index.php/RSP/article/view/8722>. Acesso em: 5 mar. 2023.

BREIMAN, L. **Random forests**. *Machine Learning*. v. 45, p. 5-32, 2001.

CHEN, D.Y. **Análise de dados com Python e Pandas**, Novarec: São Paulo, 2018

CHEN, T., GUESTRIN, C. **XGBoost: A Scalable Tree Boosting System**. Proceedings of the 22nd ACM SIGKDD Int. Conf, on Knowledge Discovery and Data Mining. São Francisco, California, USA, v.1 p.785-794, 2016.

COSTA, Eduardo Rodrigues da. **A tecnologia e estratégia na cobrança da dívida ativa da União**: Rio de Janeiro. 2020. 62 f. Dissertação (Mestrado)-FGV EBAPE - Escola Brasileira de Administração Pública e de Empresas, Rio de Janeiro, 2020.

CRUZ, Antonia Camilly Gomes. **Aprimoramento da cobrança da dívida ativa: dos nudges fiscais à conformidade tributária**. Dissertação (Mestrado, Programa de Mestrado Profissional em Direito e Gestão de Conflitos) - Universidade de Fortaleza, 2020.

FACELI, Katti; LORENA, Ana C.; GAMA, João; AL, et. **Inteligência Artificial - Uma Abordagem de Aprendizado de Máquina**. Rio de Janeiro: Grupo GEN, 2021. E-book.

FACHIN, Odete Medaua. **Fundamentos de metodologia**. 3.ed. São Paulo: Saraiva, 2001.

FÁVERO, BELFIORE, SILVA E CHAN. **Análise de dados – Modelagem Multivariada para Tomada de Decisão**. Rio de Janeiro: Editora Elsevier. 2009.

FERNANDES, A. A. T. et al. Leia este artigo se você quiser aprender regressão logística. **Revista de Sociologia e Política**, Curitiba, v.8, p.1-20, 2020.

FRIEDMAN, J. H. **Stochastic gradient boosting**. *Computational Statistics & Data Analysis*. 2002.

HAMMERSLEY, ATKINSON. P. **Ethnography: Principles in Practice**. London: Routledge, 1995.

HARRISON, M. **Machine Learning: guia de referência rápida**: trabalhando com dados estruturados em Python. São Paulo: Novatec. 2020.

IPEA. **Custo e tempo do processo de execução fiscal promovido pela Procuradoria Geral da Fazenda Nacional**. Brasília: Ipea, 2011. (Nota Técnica).

HAIR JR, Joseph F.; PRETO, William C.; BABIN, Barry J. **Análise multivariada de dados**. Porto Alegre: Grupo A, 2009. *E-book*.

CARVALHO JUNIOR, Pedro Humberto Bruno de. **O papel da inadimplência na heterogeneidade da arrecadação do IPTU no Brasil**. Brasília: Ipea, 2018.

LEITE, Leonardo Queiroz. **História da Reforma Gerencial do Estado de 1995**: São Paulo. 2019. Dissertação (Mestrado)-FGV EBAPE - Escola Brasileira de Administração Pública e de Empresas, São Paulo, 2019.

XIANGJUN Ji, TONG Weida, ZHICHAO Liu e TIELIU Shi. **Five-Feature Model for Developing the Classifier for Synergistic vs. Antagonistic Drug Combinations Built by XGBoost**. *Frontiers in Genetic* v. 10, 2019.

LI, Hua; CAO, Yumeng; LI, Siwen; ZHAO, Jianbin e Sun, YUTONG. XGBoost Model and Its Application to Personal Credit Evaluation. **International Conference on Data Mining and Big Data**. Springer, Cham: IEEE, v. 35, p. 52-61, 2020.

LUGER, George F. **Inteligência Artificial**. Trad. Daniel Vieira. Rev. Téc. Andréa Iabrudi Tavares. 6. ed. São Paulo: Pearson Education do Brasil, 2013, p. 21

MALHOTRA, Naresh K. **Pesquisa de Marketing: Uma Orientação Aplicada**. Porto Alegre: Grupo A, 2019. *E-book*.

MATIAS-PEREIRA, José. **Manual de Metodologia da Pesquisa Científica**. Rio de Janeiro: Grupo GEN, 2016. *E-book*.

MITCHELL, T M. **Machine learning**. New York: McGraw-Hill, 1997.

MUELLER, John; MASSARON, Luca. **Algoritmos Para Leigos**. São Paulo: Editora Alta Books, 2018. *E-book*.

MONARD, Maria Carolina, BARANAUSKAS, José Augusto. **Conceitos Sobre Aprendizado de Máquina. Sistemas Inteligentes Fundamentos e Aplicações**. 1 ed. Barueri-SP: Manole Ltda, 2003.

OLIVEIRA, Marcello Sartore de; SANTOS, Waldir Jorge Ladeira dos. Dívida ativa: análise da cobrança administrativa e judicial em Municípios do Rio de Janeiro. **Revista do Serviço Público (RSP)**, Brasília, p. 481-506, 2014.

OLIVEIRA, Rafael Albuquerque Gomes de. **O esgotamento do modelo judicial de cobrança de créditos da Fazenda Pública: a experiência adotada no município de Manaus a partir dos procedimentos extrajudiciais de solução de conflitos**. Dissertação (Mestrado profissional) - Fundação Getulio Vargas, Escola de Direito de São Paulo, São Paulo, 2019.

PAULA, Gilberto A. **Modelos de regressão: com apoio computacional**. Instituto de Matemática e Estatística Universidade de São Paulo. São Paulo. 2004.

RENDER, Barry; JR., Ralph M S.; HANNA, Michael E. **Análise Quantitativa para Administração**. Porto Alegre: Grupo A, 2010. *E-book*.

RIO DE JANEIRO. [Constituição (1989)]. **Constituição do Estado do Rio de Janeiro**.

RIO DE JANEIRO. **Decreto-Lei nº 05, de 15 de março de 1975**. Institui o Código Tributário do Estado do Rio de Janeiro e dá outras providências. Rio de Janeiro: Secretaria de Estado de Fazenda do Rio de Janeiro.

RIO DE JANEIRO. **Lei nº 7040, de 09 de julho de 2015**. Autoriza o estado a ceder, a título oneroso, direitos creditórios na forma que especifica, autoriza a criação de Sociedade de Propósito Específico e dá outras providências. Rio de Janeiro: Assembleia Legislativa do Rio de Janeiro.

RIO DE JANEIRO. Tribunal de Contas. **Parecer Prévio 2021**. Processo TCERJ n. 101.402-2/22. Prestação de Contas — exercício 2021. Relatório do Conselheiro Marcelo Verdini Maia.

RIBEIRO, Yasmin Arbex. **Blockchain e smart contracts: potenciais aplicações da tecnologia para a gestão da dívida ativa de um município fluminense**. 2022. 126 f. Dissertação (Mestrado em Administração Pública) - Universidade Federal Fluminense, Volta Redonda, 2022.

ROSA, A. R. A., & ALCADIPANI, R.. Modelos de gestão pública no Brasil: tradições, transformações e perspectivas. **Cadernos EBAPE**, São Paulo, v. 11, p. 556-568, 2013.

ROSSI, Gustavo Afonso Santi. **Dívida ativa tributária do município do Rio de Janeiro: um viés sobre o perfil dos sujeitos passivos inadimplentes**. 2016. 118 f. Dissertação (Mestrado em Controle de Gestão) - Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2016.

ROSSI, Gustavo Afonso Santi; SANTOS, Waldir Jorge Ladeira dos. Peculiaridades da arrecadação e cobrança da dívida ativa na administração pública. **Revista Ambiente Contábil**, Universidade Federal do Rio Grande do Norte, v.8, p. 205–225, 2016.

SANTOS, Gustavo Carvalho. **Algoritmos de Machine Learning para previsão da B3**. 2020. 90 f. Dissertação (Mestrado em Engenharia Elétrica) - Universidade Federal de Uberlândia, Uberlândia, 2020.

SÁ, Tainá Ayres. **Métodos de classificação supervisionados aplicados à identificação de fraudes de fornecedores do Estado do Rio de Janeiro**. 2022. 68 f. Dissertação (Mestrado em Ciências Contábeis) – Faculdade de Administração e Finanças, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2022.

SILVA, Denis Ribeiro da; COSTA, Daniel Fonseca da; PIMENTA, Alexandre. A Influência da Inteligência Artificial na Contabilidade e na Tributação das Organizações: uma revisão de literatura. *In: CONGRESSO USP DE CONTROLADORIA E CONTABILIDADE*, 22., 2022, São Paulo. **Anais**. São Paulo: Fipecafi, 2022. p. 1-15

SILVA, Jussara Filardi. **Transparência fiscal: uma análise das informações contábeis divulgadas pelos municípios brasileiros**. Tese (Doutorado) - Universidade de Brasília, Brasília, 2019.

SILVA, Manoella Alves da; LIMA, Diana Vaz de; FERREIRA, Lucas Oliveira Gomes. Tratamento contábil da dívida ativa tributária: um estudo de caso no GDF à luz da teoria contábil. **Revista de Contabilidade do Mestrado em Ciências Contábeis da UERJ**, Rio de Janeiro, 2016.

SCHAPIRE, R. E. **The Strenght of Weak Learnability**. Kluwer Academic Publishers, Boston. Manufactured in The Netherland. 1990.

SOUZA JUNIOR, Elio Ferreira de. **O ciclo de vida da inovação no setor público: uma análise do novo modelo de cobrança da dívida ativa do governo federal brasileiro: Rio de Janeiro**. 2020. Dissertação (Mestrado)-FGV EBAPE - Escola Brasileira de Administração Pública e de Empresas, Rio de Janeiro, 2020.

STANCHEVA-TODOROVA, Eleonora P. **How Artificial Intelligence is challenging accounting profession**. Journal of International Scientific Publications: Economy & Business, International Scientific Publications, Bulgaria, v. 12, p. 126-141, 2018.

VIRGILITO, Salvatore B. **Pesquisa de marketing: uma abordagem quantitativa e qualitativa**. Rio de Janeiro: Saraiva, 2017. *E-book*.

ANEXO – Apuração da Perda pela PGE/RJ

	Posição do Estoque da D.A. - 31/12/2022	
	Quantidade	Valor Total
Créditos tributários inscritos até 1997	5.272	R\$ 5.605.585.178,93
Créditos não tributários inscritos até 1997	105	R\$ 61.423.027,23
Total inscrito até 1997	5.377	R\$ 5.667.008.206,16
Créditos tributários inscritos após 1997	773.441	R\$ 125.511.082.656,99
Créditos não tributários inscritos após 1997	55.722	R\$ 4.615.116.524,61
Total inscrito após 1997	829.163	R\$ 130.126.199.181,60
Total Geral inscrito em Dívida Ativa	834.540	R\$ 135.793.207.387,76

Posição do Estoque da D.A. - 31/12/2022	
Valor do Ajuste p/ Perda	Estoque Recuperável
R\$ 5.587.002.531,03	R\$ 18.582.647,90
R\$ 61.367.816,71	R\$ 55.210,52
R\$ 5.648.370.347,74	R\$ 18.637.858,42
R\$ 79.795.492.469,54	R\$ 45.715.590.187,45
R\$ 4.335.432.062,37	R\$ 279.684.462,24
R\$ 84.130.924.531,91	R\$ 45.995.274.649,69
R\$ 89.779.294.879,65	R\$ 46.013.912.508,11

Tabela Dificuldade x Percentual	
Dificuldade	Faixas de Percentual
baixa	0 a 25%
média	26 a 50%
alta	51 a 100%

Créditos tributários inscritos até 1997	Quantidade	Valor Total	Dificuldade	Percentual do ajuste	Valor do Ajuste p/ Perda	Estoque Recuperável
Em parcelamento	23	R\$ 13.542.342,24	baixa	5%	R\$ 677.117,11	R\$ 12.865.225,13
Não ajuizado, valor menor	-	-	alta	100%	-	-
Com exigibilidade suspensa	27	R\$ 11.434.845,55	média	50%	R\$ 5.717.422,78	R\$ 5.717.422,78
CDA de empresas com registro de falência/concordata	825	R\$ 1.037.233.652,44	alta	100%	R\$ 1.037.233.652,44	-
Com parcelamento interrompido (PDA ou SEF)	355	R\$ 672.592.278,36	alta	100%	R\$ 672.592.278,36	-
CDA com exigibilidade plena amigável	1	R\$ 62.945,97	alta	100%	R\$ 62.945,97	-
CDA com exigibilidade plena judicial	4.041	R\$ 3.870.719.114,37	alta	100%	R\$ 3.870.719.114,37	-
Total	5.272	R\$ 5.605.585.178,93			R\$ 5.587.002.531,03	R\$ 18.582.647,90

Créditos não tributários inscritos até 1997	Quantidade	Valor Total	Dificuldade	Percentual do ajuste	Valor do ajuste p/ perda	Estoque Recuperável
Em parcelamento	1	R\$ 61.345,02	baixa	10%	R\$ 6.134,50	R\$ 55.210,52
Não ajuizado, valor menor	-	R\$ -	alta	100%	R\$ -	R\$ -
Com exigibilidade suspensa	-	R\$ -	média	50%	R\$ -	R\$ -
CDA de empresas com registro de falência/concordata	13	R\$ 46.017.166,08	alta	100%	R\$ 46.017.166,08	R\$ -
Com parcelamento interrompido (PDA ou SEF)	7	R\$ 437.784,09	alta	100%	R\$ 437.784,09	R\$ -
CDA com exigibilidade plena amigável	-	R\$ -	alta	100%	R\$ -	R\$ -
CDA com exigibilidade plena judicial	84	R\$ 14.906.732,04	alta	100%	R\$ 14.906.732,04	R\$ -
Total	105	R\$ 61.423.027,23			R\$ 61.367.816,71	R\$ 55.210,52

Créditos tributários inscritos após 1997	Quantidade	Valor Total	Dificuldade	Percentual do ajuste	Valor do ajuste p/ perda	Estoque Recuperável
Em parcelamento	19.259	R\$ 1.747.178.666,23	baixa	5%	R\$ 87.358.933,31	R\$ 1.659.819.732,92
Não ajuizados, valor menor	549.443	R\$ 932.459.823,42	alta	100%	R\$ 932.459.823,42	R\$ -
Com exigibilidade suspensa	2.392	R\$ 6.930.697.347,09	média	50%	R\$ 3.465.348.673,55	R\$ 3.465.348.673,55
CDA de empresas com registro de falência/concordata	13.220	R\$ 18.225.879.801,93	alta	100%	R\$ 18.225.879.801,93	R\$ -
Com parcelamento interrompido (PDA ou SEF)	26.398	R\$ 10.054.261.944,64	alta	100%	R\$ 10.054.261.944,64	R\$ -
CDA com exigibilidade plena amigável	62.981	R\$ 14.027.112.610,04	média	50%	R\$ 7.013.556.305,02	R\$ 7.013.556.305,02
Ajuizados há menos de 10 anos com exigibilidade plena	93.371	R\$ 67.153.730.951,93	média	50%	R\$ 33.576.865.475,97	R\$ 33.576.865.475,97
Ajuizados há 10 anos ou mais com exigibilidade plena	6.377	R\$ 6.439.761.511,71	alta	100%	R\$ 6.439.761.511,71	R\$ -
Total	773.441	R\$125.511.082.656,99			R\$ 79.795.492.469,54	R\$ 45.715.590.187,45

Créditos não tributários inscritos após 1997	Quantidade	Valor Total	Dificuldade	Percentual do ajuste	Valor do ajuste p/perda	Estoque Recuperável
Em parcelamento	743	R\$ 15.175.508,07	baixa	10%	R\$ 1.517.550,81	R\$ 13.657.957,26
Não ajuizado, valor menor	23.408	R\$ 100.271.258,03	alta	100%	R\$ 100.271.258,03	R\$ -
Com exigibilidade suspensa	349	R\$ 423.941.036,79	média	50%	R\$ 211.970.518,40	R\$ 211.970.518,40
CDA de empresas com registro de falência/concordata	1.995	R\$ 135.092.874,70	alta	100%	R\$ 135.092.874,70	R\$ -
Com parcelamento interrompido (PDA ou SEF)	1.757	R\$ 106.092.137,71	alta	100%	R\$ 106.092.137,71	R\$ -
CDA com exigibilidade plena amigável	355	R\$ 108.111.973,17	média	50%	R\$ 54.055.986,59	R\$ 54.055.986,59
Ajuizados há menos de 10 anos com exigibilidade plena	26.850	R\$ 3.644.585.502,82	alta	100%	R\$ 3.644.585.502,82	R\$ -
Ajuizados há 10 anos ou mais com exigibilidade plena	265	R\$ 81.846.233,32	alta	100%	R\$ 81.846.233,32	R\$ -
Total	55.722	R\$ 4.615.116.524,61			R\$ 4.335.432.062,37	R\$ 279.684.462,24

Créditos não tributários inscritos após 1997	Quantidade	Valor Total	Dificuldade	Percentual do ajuste	Valor do ajuste p/ perda	Estoque Recuperável
Em parcelamento	743	R\$ 15.175.508,07	baixa	10%	R\$ 1.517.550,81	R\$ 13.657.957,26
Não ajuizado, valor menor	23.408	R\$ 100.271.258,03	alta	100%	R\$ 100.271.258,03	R\$ -
Com exigibilidade suspensa	349	R\$ 423.941.036,79	média	50%	R\$ 211.970.518,40	R\$ 211.970.518,40
CDA de empresas com registro de falência/concordata	1.995	R\$ 135.092.874,70	alta	100%	R\$ 135.092.874,70	R\$ -
Com parcelamento interrompido (PDA ou SEF)	1.757	R\$ 106.092.137,71	alta	100%	R\$ 106.092.137,71	R\$ -
CDA com exigibilidade plena amigável	355	R\$ 108.111.973,17	média	50%	R\$ 54.055.986,59	R\$ 54.055.986,59
Ajuizados há menos de 10 anos com exigibilidade plena	26.850	R\$ 3.644.585.502,82	alta	100%	R\$ 3.644.585.502,82	R\$ -
Ajuizados há 10 anos ou mais com exigibilidade plena	265	R\$ 81.846.233,32	alta	100%	R\$ 81.846.233,32	R\$ -
Total	55.722	R\$ 4.615.116.524,61			R\$ 4.335.432.062,37	R\$ 279.684.462,24