

Impacto da Inteligência Artificial na Atividade de Auditoria: Equacionando Gargalos nos Repasses da União para Entes Subnacionais

Sérgio Tadeu Neiva Carvalho

RESUMO

Objetivos: Esta pesquisa visa compreender como o uso da Inteligência Artificial (IA) é suficientemente percuciente para dispensar a análise convencional de prestação de contas de convênios. A utilização de soluções tecnológicas como *Machine Learning* (ML) e IA vem se ampliando no setor privado com impactos positivos na prevenção às fraudes e no incremento da eficiência. Entretanto, os estudos acerca de seu uso na Administração Pública ainda são incipientes. Logo, este trabalho almeja, também, compreender os impactos da IA no Governo. A metodologia valeu-se da implementação de um ambiente computacional para testar arquiteturas de ML, parametrizando-as conforme a volumetria e a identidade do órgão concedente. Os resultados apontam que a identidade do órgão concedente influencia o desempenho do ML e à medida que o tamanho da sequência de treinamento cresce, o desempenho aumenta. A compreensão do desempenho da IA validará essa abordagem inovadora, racionalizando a força de trabalho no setor público.

Palavras-Chave: *Controle, Federalismo, Inteligência Artificial, Administração Pública, Transferências Discricionárias*

Classificação JEL: H83, H77, M42, P35

SUMÁRIO

SUMÁRIO.....	2
1. INTRODUÇÃO.....	3
2. REFERENCIAL TEÓRICO.....	6
2.1. Relações Federativas e Transferências Voluntárias.....	6
2.2. O papel do uso da tecnologia aplicada ao Controle.....	17
2.3. A Sofisticação dos métodos de Auditoria: Técnicas de mensuração de desempenho de <i>Machine Learning</i>	20
3. CONTEXTUALIZAÇÃO DO PROBLEMA.....	25
3.1. Transferências Voluntárias e a Plataforma +Brasil (antigo SICONV).....	25
3.2. O Sistema “ <i>Malha Fina de Convênios</i> ”.....	28
3.3. Descrição do problema.....	32
3.4. Importância do Problema.....	36
4. METODOLOGIA.....	39
4.1. Relevância para o algoritmo de IA em distinguir a identidade dos repassadores de recursos.....	45
4.2. Relevância do tamanho da série de treinamento para o algoritmo IA.....	48
5. ANÁLISE DOS RESULTADOS.....	52
5.1.1. Métricas <i>AUC</i> e <i>inexatidão</i> (ϵ) conforme a identidade do órgão concedente.....	52
5.1.1.1. Cenário 1: algoritmo de IA treinado com o universo e aplicado em cada órgão concedente separadamente.....	52
5.1.1.2. Cenário 2: algoritmo de IA treinado com convênios conforme a identidade do órgão concedente.....	56
5.1.2. Métricas <i>AUC</i> e <i>inexatidão</i> (ϵ) conforme a volumetria da sequência de treinamento.....	62
6. CONSIDERAÇÕES FINAIS.....	66
REFERÊNCIAS.....	70
ANEXO I - CÓDIGO PHYTON UTILIZADO NAS SIMULAÇÕES.....	79

1. INTRODUÇÃO

Este trabalho tem como objetivo analisar como a integração de tecnologia com a auditoria interna pode ser útil no controle das Transferências Voluntárias da União, intermediadas pela Plataforma +Brasil¹. Nessa perspectiva, analisa-se o caso do sistema “*Malha Fina de Convênios*”, abordagem tecnológica inovadora envolvendo a Controladoria-Geral da União – CGU - e o Ministério da Economia - ME. Nesse bojo emerge a relevância de dois pontos centrais: o papel do Controle e a aplicação da Tecnologia para o seu desempenho.

O uso da Tecnologia revela sua importância porque considera a natureza variável dos registros e a incorporação de fontes não tradicionais de dados nos domínios da auditoria (Vasarhelyi, Kogan e Tuttle 2015). As novas oportunidades de análise de auditoria proporcionada pelo uso da Tecnologia vão ao encontro da necessidade de alterações de controle. Contudo, existem desafios no caminho de sua utilização, além dos seus benefícios potenciais.

O percurso de superação desses obstáculos na Administração Pública possui nuances intrínsecos. Wirtz, Weyerer e Geyer (2019) propõem 4 categorias diferentes de desafios a serem superados, especificamente para o setor público: (i) implementação da tecnologia, (ii) legislação e regulação da aplicabilidade da Tecnologia, (iii) ética e (iv) paradigma comportamental. Por sua vez, o paradigma comportamental enseja um debate mais complexo em relação às outras categorias, haja vista que aborda temas tortuosos de uma burocracia (Barzelay 2019; Bresser-Pereira 2008; Mintzberg 1979; Olivieri 2008), como a substituição e a transformação da força de trabalho, a aceitação e a confiança da adoção de tecnologias em detrimento dos métodos convencionais de trabalho, e, por último, a interação homem-máquina. Issa, Sun e Vasarhelyi (2016) propõem uma reflexão sobre os efeitos que a automação gera na substituição de auditores como força de trabalho.

Entretanto, a adoção da Tecnologia na atividade de auditoria deve ser encorajada. Uma vez que o auditor aprende a lidar apropriadamente com as técnicas de mineração de dados, inteligência artificial, Lei de Benford, entre outras, há uma esclarecedora ruptura dos métodos vigentes, levando-o a questionar como as auditorias eram feitas sem o emprego dessas modernidades tecnológicas (Huang e Vasarhelyi 2019; C. Zhang 2019). Notadamente, a

¹ O Sistema de Convênios do Governo Federal – SICONV – foi rebatizado para Plataforma +Brasil (Decreto n. 10.035, 2019).

Administração Pública precisa buscar incessantemente a melhoria no desempenho do Controle para que as políticas públicas sejam avaliadas e monitoradas, com menor custo e com mais rapidez. O uso de inovações tecnológicas na função Controle permite o avanço nessa seara (Maciejewski 2017).

Enquanto, por um lado, a Tecnologia aplicada ao Controle é relevante para a execução das Políticas Públicas, por outro lado, a implementação delas no Brasil é um enorme desafio. Somos um país com dimensões continentais, diversidades culturais profundas, configurados em um número suntuoso de entidades subnacionais independentes, contemplando 27 unidades da federação e 5570 municípios (IBGE, 2020). Por sua vez, as Transferências Voluntárias da União são uma opção útil para a implantação de políticas públicas no país.

As Transferências Voluntárias são recursos repassados pela União, sem determinação constitucional ou legal, a outros entes federados ou a entidades privadas sem fins lucrativos para a realização de obras ou serviços de interesse comum. O Estado federal brasileiro utiliza o federalismo por cooperação para a consecução dos princípios e objetivos fundamentais definidos na Constituição Federal de 1988 - CF88 -, artigos 1º e 3º, em especial o desenvolvimento nacional e a redução das desigualdades sociais e regionais.

Todavia, o arranjo institucional estipulado pela CF88 atribui competências concorrentes entre os entes, nem sempre harmoniosas, na medida em que a estrutura federativa é muito desigual entre as esferas municipais, estaduais e federal (Fajardo, 2016). Com efeito, a realização de políticas públicas pelo Governo Federal por meio de parcerias e convênios com os entes subnacionais é imperiosa, mas nem sempre os repasses de recursos são redistributivos, conforme demonstram as pesquisas empreendidas por Amorim Neto e Simonassi (2013), Brollo e Nannicini (2012), Ferreira e Bugarin (2008), Meireles (2019), e Soares e Melo (2016).

Por sua vez, as avaliações da política pública das Transferências Voluntárias da União indicam um flagrante desequilíbrio entre a quantidade de convênios celebrados e a capacidade dos órgãos concedentes analisarem as prestações de contas das entidades convenentes (Brasil, 2018a). A expressão dessa realidade é o acúmulo contínuo de prestação de contas pendente de análise. A consequência prática desse fenômeno é a existência de um estoque incremental de passivo de convênios, cuja prestação de contas não possui perspectiva algum de ser analisada tempestivamente, afastando o preceito constitucional do dever de prestar contas. Outrossim, as exigências impostas pelos órgãos de controle, notadamente a CGU, aos órgãos concedentes e convenentes quanto aos requisitos de comprovações de despesas contribuem para a morosidade

do processo de prestação de contas. Logo, a CGU também é parte do problema, embora indiretamente.

Diante desse cenário paradoxal, em que a atuação do órgão de controle implica na deficiência da política pública que ele mesmo controla, a CGU fomentou a propositura de uma solução por meio do uso da tecnologia nas ações de controle. Essa solução, denominada “*Malha Fina de Convênios*”, facilita a análise de prestação de contas de convênios, observando o expressivo volume de instrumentos celebrados e a limitada capacidade de análise do corpo de servidores dos órgãos concedentes. O sistema “*Malha Fina de Convênios*” é um modelo preditivo de Inteligência Artificial que permite indicar, com determinado grau de certeza, o resultado da análise das prestações de contas dos convênios no momento em que suas contas são apresentadas pelos convenientes (entes subnacionais recebedores de recursos) aos concedentes (órgãos da União repassadores de recursos).

Nesse contexto, as potencialidades do uso do Malha Fina de Convênios, seus impactos já demonstrados e suas limitações ainda não são claros e definitivos. Dessa forma, o objetivo do trabalho é analisar alguns desses aspectos, especificamente no que tange a precisão e a acurácia do Sistema Malha Fina de Convênios em diferentes circunstâncias. Analisar-se-á a influência do tamanho e a identidade dos órgãos repassadores de recursos, em diferentes cenários de sequência de treinamento, na acurácia e na precisão dos algoritmos (Alpaydin 2014; Breiman 2001; Domingos 2012). A implementação desses diferentes cenários será conduzida por meio da utilização de aprendizado de máquina supervisionado *Random Forest*, executado em linguagem Phyton.

Para entender o impacto do “*Malha Fina de Convênios*” no mundo real das Transferências Voluntárias da União, este estudo se centra no tema das relações entre auditoria e tecnologia (Issa, Sun, e Vasarhelyi 2016), questão clássica para a função Controle na ciência Administração contemporânea, uma vez que a forte expansão da delegação de políticas públicas aos entes subnacionais a partir da Constituição Federal de 1988 ampliou a relevância da conexão entre incremento de recursos transferidos pela União e a necessidade de fiscalizá-los eficientemente (Arretche 2010; Abrucio e Franzese 2007; Afonso, Araújo, e Fajardo 2016). Tal questão será tratada tanto do ponto de vista teórico como empírico, neste caso, analisando-se a aplicabilidade do “*Malha Fina de Convênios*” no Governo Federal.

O presente estudo analisará essa temática em 5 etapas, além desta Introdução. A seção 2, subsequente a esta seção introdutória, consiste em uma revisão da literatura, iniciando-se

com uma breve visão geral das Transferências Voluntárias da União, dos conceitos do Controle e sua importância na Administração Pública, além de destacar o papel da tecnologia no redesenho da atividade de auditoria, abordando conceitos de *machine learning*.

Em seguida, a seção 3 aborda o problema da prestação de contas dos repasses da União operacionalizados na Plataforma +Brasil. Ainda nessa seção é exposto o sistema “*Malha Fina de Convênios*” como proposta de solução para o passivo de contas a serem analisadas, bem como a caracterização de suas limitações e potencialidades. Por sua vez, o problema de pesquisa é delineado e as hipóteses a serem testadas são elencadas. Posteriormente, na seção 4 são apresentados os procedimentos metodológicos adotados, descrevendo a estratégia de testes de simulação do aprendizado de máquina e o detalhamento da base de dados da Plataforma +Brasil, incluindo o seu corte temporal. Ato contínuo, na seção 5 são apresentados os resultados obtidos nas simulações e os dados empíricos. As análises estão apresentadas conforme duas arquiteturas configuradas para simular o desempenho dos algoritmos de aprendizado de máquina: identidade do repassador de recursos e volumetria da sequência de treinamento. Por fim, na seção 6, apresentam-se as considerações finais, explicitando-se as principais conclusões, bem como as limitações do presente estudo e algumas propostas de trabalho para futuras pesquisas.

2.REFERENCIAL TEÓRICO

Esta seção aborda o referencial teórico utilizado para o embasamento do estudo, ou seja, busca-se aqui revisar a literatura já desenvolvida, ressaltando as potenciais contribuições deste trabalho. Feitas as considerações sobre as relações federativas e sua influência nas Transferências Voluntárias, passa-se à discussão do Controle no arcabouço racional-legal da Administração Pública e, em seguida, à contextualização do papel da Tecnologia aplicada ao Controle no paradigma da “*Digital Era Governament*”. Dessa forma, o objetivo desta seção é remeter a aplicabilidade de uma nova tecnologia para o uso do Controle com o propósito de solucionar um problema real de uma política pública, qual seja as Transferências Voluntárias.

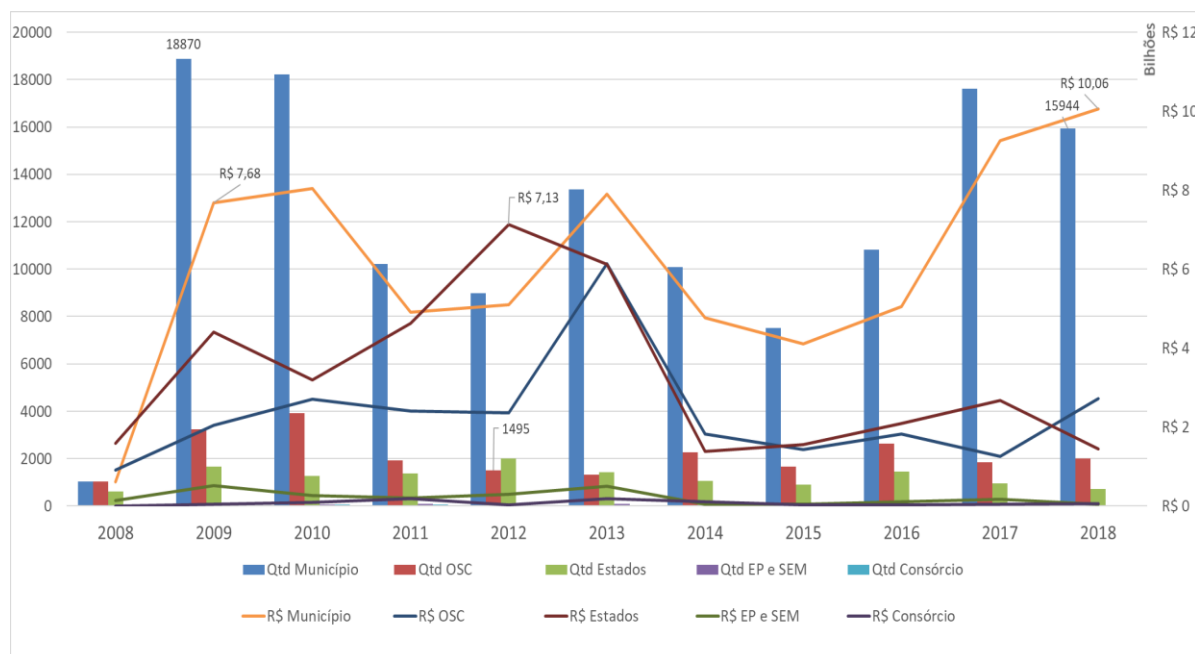
2.1. Relações Federativas e Transferências Voluntárias

As Transferências Voluntárias da União são uma política pública em que a União repassa recursos públicos federais discricionariamente a Estados, Municípios, Distrito Federal ou entidades privadas sem fins lucrativos. O propósito para o repasse de recursos federais aos entes

federados ou às organizações sociais serve para que esses executem políticas em áreas de educação, saúde, saneamento, construção e recuperação de estradas, abastecimento de água, habitação e energia urbana e rural. Afinal de contas, o cidadão mora no município (Amorim Neto e Simonassi 2013) e a execução mais assertiva da política pública só poderia ocorrer nesse ente subnacional (Soares e Melo 2016).

As transferências voluntárias desempenham um importante papel na relações federativas brasileiras, principalmente entre União e municípios. Segundo Fajardo (2016), a Constituição Federal de 1988 (CF88) foi um marco no arranjo institucional da federação brasileira uma vez que conferiu aos municípios a categoria de ente federativo, revestindo-os de uma importância singular, tendo em vista sua representatividade maior em termos de quantidade em relação aos Estados. Por sua vez, Arretche (2010) e outros acadêmicos, como Afonso, Araújo e Fajardo (2016) e Loureiro e Abrucio (2004), destacam a peculiaridade do advento do ente municipal no modelo do federalismo brasileiro, cada um em sua perspectiva decorrente de suas pesquisas. Em resumo, grande parte dessas perspectivas são negativas porque as evidências empíricas desses estudos mostram que arrecadação pequena desses entes subnacionais e sua grande parcela na composição do gasto público ameaçam a sustentabilidade do ente municipal. Igualmente, essas perspectivas negativas são corroboradas com a assertiva de Amorim Neto e Simonassi (2013) em que a captação de apoio político prepondera sobre critérios de equidade nas transferências da União para os entes subnacionais. Não obstante, o município exerce um papel central nas transferências voluntárias da União. De certa forma, o protagonismo dos municípios na política pública das transferências voluntárias da união pode ser observado na Figura 1, na medida em que é o destinatário com maior volume e materialidade de convênios, ultrapassando até mesmo as entidades privadas sem fins lucrativos.

Figura 1 – Distribuição do quantitativo e materialidade de instrumentos operacionalizados na Plataforma +Brasil.



Fonte: Plataforma +Brasil. Elaborado pelo autor².

Afora a flagrante supremacia dos municípios em relação aos estados em termos de quantidade de instrumentos, a Figura 1 ilustra um panorama menos expressivo em relação a materialidade de repasse feito aos Municípios em relação aos Estados, ainda que esses sejam superiores ao primeiro destinatário quando comparado ao segundo. A Figura 1 representa as quantidades de instrumentos pelas barras e os valores de recursos transferidos pelas linhas. Com efeito, as barras azuis, que simbolizam a quantidade de instrumentos em que os Municípios configuram como beneficiados, são muito maiores em magnitude do que qualquer outra categoria de barra, ou seja, os Estados e as entidades sem fins lucrativos. Entretanto, a linha laranja, que simboliza o valor total de recursos transferidos aos Municípios, não está tão distante das linhas que simbolizam os valores repassados aos Estados e às entidades sem fins lucrativos, destacando-se que em 2012 o volume de aporte feito aos Estados superou àquele feito aos municípios. Todavia, a Tabela 1 permite inferir que o patamar de recursos transferidos aos Municípios é quase o dobro daqueles transferidos aos Estados.

² EP- Empresa Pública, SEM – Sociedade de Economia Mista, OSC – Organização da Sociedade Civil.

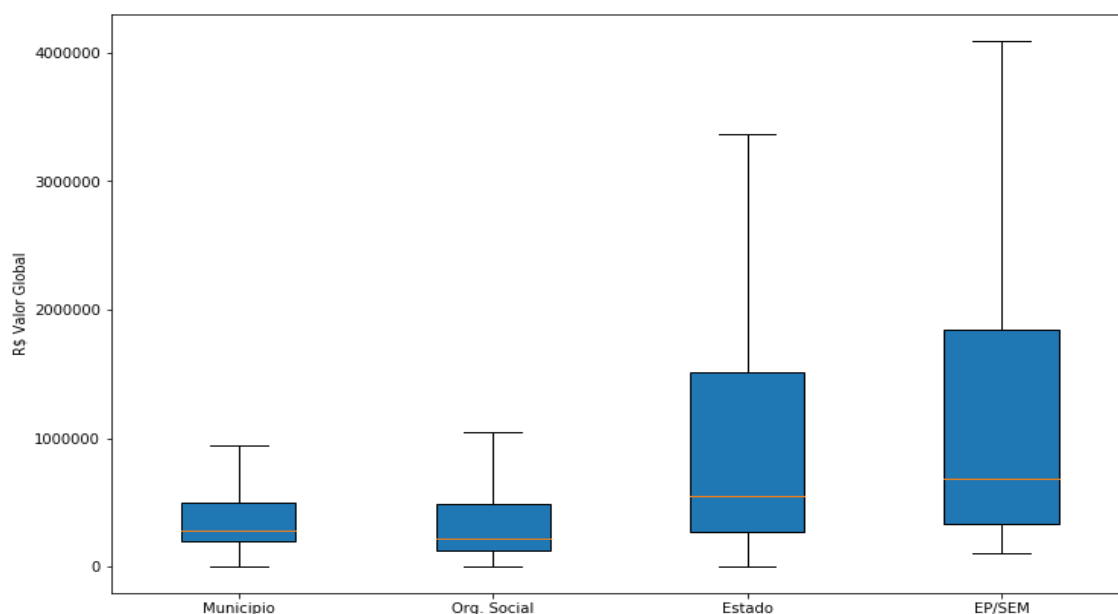
Tabela 1 - Quantitativo e Valores Totais de instrumentos operacionalizados na Plataforma +Brasil por natureza jurídica do destinatário

Natureza Jurídica	Quantidade	Valor
Município	136.429	R\$ 69.319.120.959,84
Organização da Sociedade Civil	24.118	R\$ 26.614.140.309,16
Estados	13.624	R\$ 36.598.277.602,35
Empresas Públicas e Sociedade de Economia Mista	852	R\$ 2.331.638.055,77
Consórcio Público	433	R\$ 815.383.238,07
Total	175456	R\$ 135.678.560.165,19

. Elaborado pelo autor. Fonte: Plataforma +Brasil.

Em certa medida, a Figura 1 e a Tabela 1 robustecem o protagonismo que o Município exerce na política pública das transferências voluntárias da União e conseqüentemente no arranjo federativo (Arretche 2010). Ademais, o cenário apresentado permite constatar que os instrumentos celebrados com os municípios são quantitativamente volumosos e possuem valores baixos, conforme observado na Figura 2. Isso contribui para a pulverização de recursos, o que pode gerar efeitos adversos, haja vista que nem sempre há coordenação para execução de programas por parte do governo federal e os governos subnacionais adotam posturas exageradamente autônomas (Abrucio 2005; Fajardo 2016). Naturalmente, políticas públicas são sombreadas, a grande maioria dos municípios brasileiros não contam com um estrutura burocrática adequada para o gerenciamento de grande quantidade de convênios, ocorre sobrecarga na capacidade operacional para gerenciá-los nos órgãos da União e, principalmente, no município, e por último e mais impactante, o fracionamento dos recursos transferidos materializando-se em convênios de baixo valor (Figura 2) faz com que a efetividade de programas de governo seja baixa, contrapondo com a hipótese de os recursos serem aplicados de maneira centralizada, em grande escala. Igualmente, os custos envolvidos para gerenciar o ciclo de vida de convênios de alto valor não é diferente do custo daqueles de baixo valor. Esses desafios, por vezes, afetam a boa aplicação dos recursos públicos oriundos de transferências voluntárias.

Figura 2 - Boxplot dos Valores Globais dos convênios segregado por Ente.



Elaborado pelo Autor. Fonte: Plataforma +Brasil.

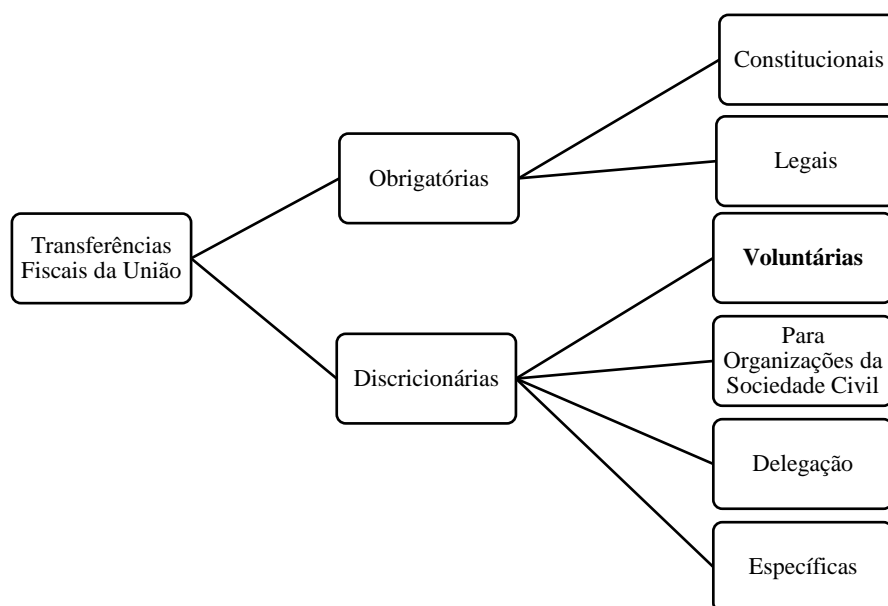
A constituinte de 1988 adotou uma fórmula federativa que previu a cooperação intergovernamental vertical, sendo que as transferências voluntárias são uma ferramenta para a consecução desse objetivo. A criação dos municípios após a promulgação da CF88 vai ao encontro desse preceito e irradia efeitos sob o prisma político, o que remete a existência de uma articulação concebida ao longo do tempo, pretérita ao evento de promulgação da Carta Constitucional. Fajardo (2016), por sua vez, argumenta, ancorado nas pesquisas promovidas por Abrucio (2005) e Arretche (2010), que a CF88 foi um processo que se iniciara bem antes da sua promulgação, promovendo consequências sob os primas administrativos e fiscais, além do ponto de vista político, cujo maior impacto foi a criação do ente municipal.

Entretanto, a pluralidade de temas com competências concorrentes entre os entes federados e a ampla diversidade de legislação correlata, abrangendo desde saúde, perpassando por educação, segurança, cultura, até a assistência social, torna muito difícil a existência de uma classificação precisa dos tipos de transferências da União, implicando em divergências na

doutrina e academia (Brasil, 2008; Dallaverde, 2014)). Conquanto a existência dessas intempéries, órgãos decisores das políticas de relacionamento federativo e gerenciamento da despesa pública, situados em posição central no Governo Federal, empreenderam esforços para elucidar o panorama das transferências fiscais da União. Constituiu-se um grupo formado pelo extinto Ministério do Planejamento (SOF e SLTI), Ministério da Fazenda (STN) e CGU, liderado pela STN, cujo resultado do trabalho foi a categorização das transferências fiscais da União, conforme disposto na Nota Técnica nº 14/2015/COINT/SURIN/STN/MF-DF.

As transferências da União admitem, segundo a nota supracitada, as classificações (i) *obrigatórias* e (ii) *discricionárias*, conforme ilustrado na Figura 3. As transferências obrigatórias compreendem aquelas que possuem previsão expressa na Constituição Federal ou Lei e são transferidas diretamente, uma que vez que não há necessidade de celebração de instrumento jurídico formal entre a parte concedente e a beneficiada. Ato contínuo, as transferências emanadas diretamente do ordenamento da Constituição subclassificam-se como *constitucionais*, enquanto aquelas em que a obrigatoriedade tem sede em Lei subclassificam-se como *legais*. São exemplos de transferências *obrigatórias constitucionais* aquelas realizadas para o Fundo de Participação dos Municípios – FPM - ou Fundo de Participação dos Estados – FPE. Por outro lado, as transferências *obrigatórias legais* podem ser encontradas em políticas públicas de educação básica como nos PNAE, PNAT, PDDE, tanto como nas políticas de saúde no Sistema Único de Saúde – SUS - e no Fundo Nacional de Saúde-FNS.

Figura 3 – Classificação da STN para os tipos de transferências fiscais da União



Elaborado pelo Autor.

Por seu turno, as transferências *discricionárias* compreendem os recursos repassados sob a égide de um instrumento jurídico celebrado entre duas partes e subclassificam-se em 4 categorias. A primeira subclassificação, e mais importante para o propósito deste estudo, são as transferências *discricionárias voluntárias*, consistindo naquelas em que a entrega de recursos aos entes subnacionais almeja a cooperação, auxílio ou assistência financeira. As transferências *discricionárias voluntárias* não são destinadas ao SUS, como também não estão enquadradas nas hipóteses das transferências *obrigatórias*. Comumente denominadas como transferências voluntárias no cotidiano da burocracia federal, as transferências *discricionárias voluntárias* são operacionalizadas na Plataforma +Brasil e consistem basicamente em dois tipos de instrumentos jurídicos: o convênio e o contrato de repasse. A diferença entre esses instrumentos reside no fato de que o primeiro é celebrado diretamente entre as partes, de um lado um órgão do ente federal e de outro um órgão do ente subnacional, e o segundo é celebrado por intermédio de uma agente financeiro federal no papel de mandatária da União. Atualmente, esse papel é exercido pela CAIXA. Geralmente, os objetos de convênios têm escopo para aquisição de bens e contratação de serviços enquanto os objetos de contratos de repasse consistem em obras de construção civil.

A importância ostentada pelas transferências voluntárias (transferências *discricionárias voluntárias*) no arranjo institucional federal do estado brasileiro não está refletida em sua

legislação ordinária. Com efeito, o volume de transferências voluntárias é expressivamente inferior ao volume da categoria transferências *obrigatórias*. Em 2018 as transferências voluntárias totalizaram mais de R\$ 14 bilhões³, enquanto as transferências obrigatórias atingiram patamar superior a R\$ 350 bilhões⁴. Porém, isso não é motivo para rotular aquelas como insignificantes nem tampouco inexpressivas em relação a estas. Pelo contrário, os recursos se mostram altamente expressivos para os entes que dependem de recursos federais para a realização de novos investimentos (Fajardo 2016), tendo em vista que as receitas de grande parte dos Municípios, sejam elas de sua própria arrecadação ou transferidas por força de determinação constitucional, acabam sendo quase que integralmente empregadas nas despesas correntes para manutenção da máquina administrativa (Abrucio 2005). Ocorre que as transferências *obrigatórias* possuem respaldo normativo robusto, enquanto as transferências *discricionárias voluntárias* destacam-se pela ausência de legislação ordinária detalhada.

O arcabouço normativo das transferências voluntárias da União emana da expressa previsão legal no artigo 116 da Lei Federal n. 8.666/93. Existem estudos indicando que as transferências voluntárias estão amparadas no dispositivo constitucional (art. 241), entre eles Dallaverde (2016), embora essa linha de pensamento é silente na Doutrina do Direito Administrativo brasileiro, notadamente em Pietro (2015):

“(…)

*Art. 241. A União, os Estados, o Distrito Federal e os Municípios disciplinarão por meio de lei os consórcios públicos e os **convênios de cooperação entre os entes federados**, autorizando a gestão associada de serviços públicos, bem como a transferência total ou parcial de encargos, serviços, pessoal e bens essenciais à continuidade dos serviços transferidos.*

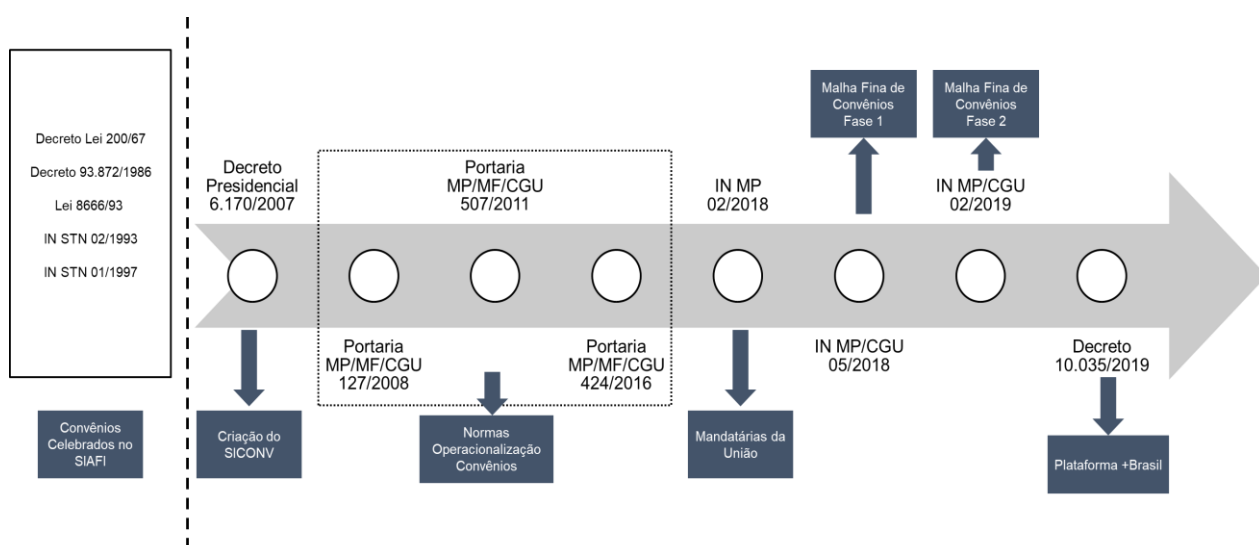
(…)” **grifo meu.**

3 Dados obtidos na Plataforma +Brasil. Trata-se da soma dos valores globais de todos os instrumentos cadastrados no sistema. Os valores globais são a parcela repassada pela União e contrapartida aportada pelos convenientes. O valor total de convênios celebrados em 2018 atingiu R\$ 14.335.621.107,03.

4 Dados obtidos no Sistema Integrado de Administração Financeira – SIAFI. As transferências obrigatórias incluem Fundo de Participação dos Estados e do Distrito Federal - FPE; Fundo de Participação dos Municípios - FPM; IPI - Exportação; CIDE-Combustíveis; Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação - Fundeb; Royalties; Imposto sobre a Propriedade Territorial Rural – ITR; IOF-Ouro; e Lei Kandir (LC 87/96). O valor total transferido para estados e municípios em 2018 atingiu R\$ 351.447.109.424,57.

De certa forma, o desamparo de diplomas legais regulando a política pública de transferências voluntárias é um reflexo da contemporaneidade recente da alçada do município para ente federado, inovação proporcionada pela CF88. As transferências voluntárias, por se tratar de uma política discricionária com aplicabilidade em amplo espectro de temas, não incentiva os atores do processo em discipliná-lo mediante legislação aprovada no Congresso Nacional, relegando-os para normas infralegais (Pietro 2015). Nesse quesito, a lacuna legal foi suplantada com uma miríade de portarias e decretos editadas ao longo do tempo, conforme se observa na Figura 4. Pode-se dizer que as transferências voluntárias ganham contornos radicalmente diferentes após a criação do SICONV em 2007 (Meireles 2019), na medida em que todo o processo de gerenciamento dos convênios com o governo federal passou a ser feito nesse sistema. Anteriormente, os convênios celebrados com a União eram registrados no SIAFI, mas esse sistema carecia de funcionalidades para permitir o controle assertivo das transferências, prejudicando a transparência de informações e a divulgação de dados, e, por outro lado, favorecendo o cenário de prestações de contas pendentes e a malversação dos recursos públicos. Ora, o SIAFI foi desenvolvido para servir como um sistema de execução financeira do Governo Federal (Abrucio e Loureiro 2018) e não um sistema de gerenciamento de convênios.

Figura 4 - Evolução do arcabouço normativo das transferências voluntárias.

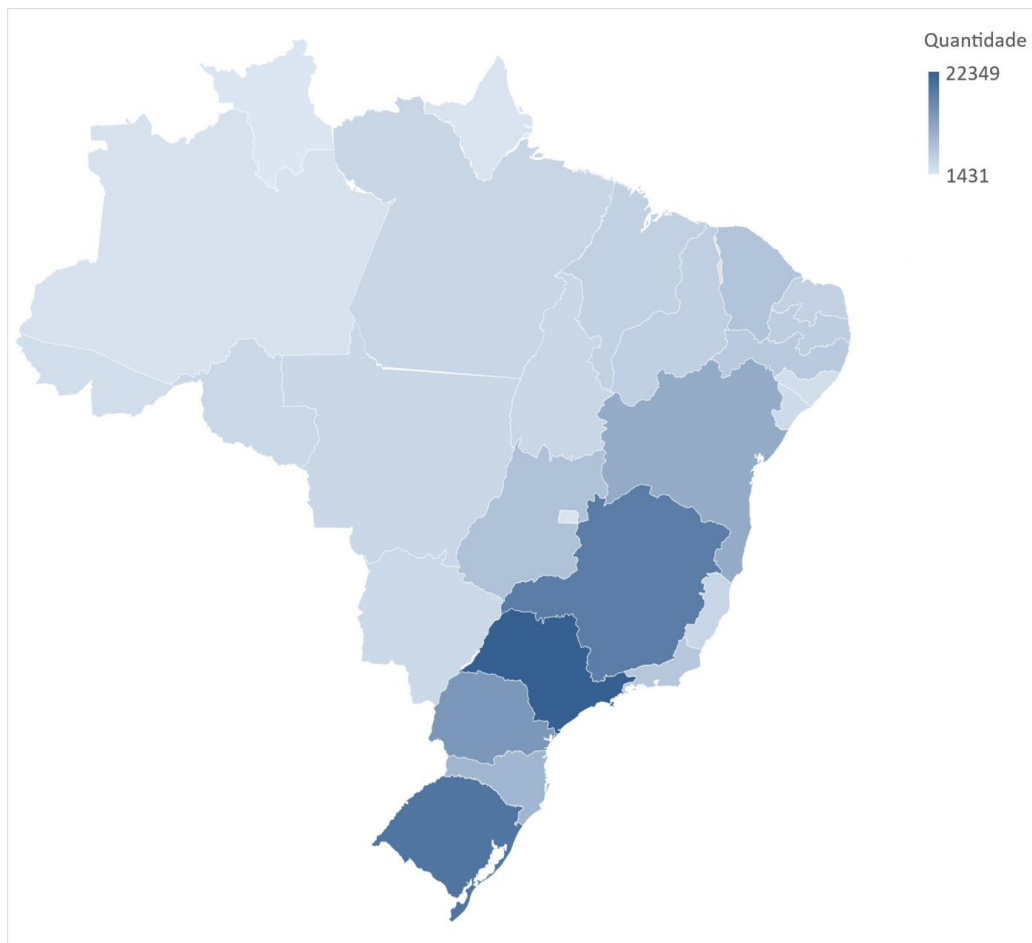


Elaborado pelo Autor.

A Instrução Normativa nº 01/97 da STN se destacou como um importante marco regulamentador das transferências voluntárias, tendo perdurado até a edição do Decreto n. 6.170 de 2007, o qual revolucionou o processo com a criação do SICONV. Em 2008, a Portaria Interministerial 127 foi editada com o propósito de veicular as normas operacionais para a execução do Decreto n. 6.170 (2007) e foi sucedida pelas Portarias Interministeriais 507 e 424 de 2011 e 2016, respectivamente. Em seguida, a cronologia apresentada na Figura 4 evidencia outro grande marco normativo, a Instrução Normativa n. 2 de 2018, que tratou de estabelecer diretrizes e regras para a execução de contratos de repasse (tipo de instrumento das transferências *discricionárias voluntárias*) por intermédio das Mandatárias da União. Ainda em 2018, houve a implementação da primeira fase do sistema “*Malha Fina de Convênios*”, com a edição da Instrução Normativa CGU/MP n. 5 de 2018, viabilizando uma alternativa automatizada para a análise de prestação de contas. Nessa medida, em 2019 iniciou-se a segunda fase do “*Malha Fina de Convênios*”, por meio da publicação da Instrução Normativa CGU/MP n. 2 de 2019. Enquanto a primeira fase restringiu o escopo da análise automatizada apenas ao estoque de passivo de prestações de contas pendentes de análise, a segunda fase permitiu o uso da análise automatizada em prestações de contas anacronicamente, conquanto o estoque fosse saneado.

As considerações abordadas em epígrafe ilustram que as transferências voluntárias da união não figuram com uma política pública isolada, muito menos um fim em si mesma, haja vista que sua finalidade primordial é possibilitar, a título de cooperação, o financiamento dos serviços públicos nos entes subnacionais. Nesse ponto, pode surgir um questionamento se a cooperação financeira recebida por um ente mais pobre, em detrimento de sua pequena contribuição comparada a um ente mais rico, contraria o equilíbrio do pacto federativo. Em resposta, consagra-se a modulação que os desígnios das relações federativas aplicam nas transferências voluntárias, pois o equilíbrio federativo solidifica-se com o mecanismo de repasse de recursos (Afonso, Araújo, e Fajardo 2016; Amorim Neto e Simonassi 2013; Fajardo 2016). Objetiva-se a inclusão dos entes mais carentes e necessitados no processo de desenvolvimento, e não a exclusão dos entes mais favorecidos, responsáveis muitas vezes por impulsionar a economia da Federação.

Figura 5 – Concentração de Instrumentos Celebrados por Unidade da Federação.



Elaborado pelo Autor. Fonte: Plataforma +Brasil.

Por derradeiro, é imprescindível que os critérios de equidade se sobreponham aos objetivos políticos nos repasses decorrentes de transferências voluntárias, pois assim esse recurso contribuiria para a diminuição das desigualdades inter-regionais na federação brasileira. Entretanto, as evidências empíricas das pesquisas de Amorim Neto e Simonassi (2013), Brollo e Nannicini (2012), Limongi e Figueiredo (2005), Meireles (2019) e Soares e Melo (2016) indicam o contrário, retratando que os entes subnacionais com alinhamento político com o mandatário da União ou com maior representatividade no Congresso Nacional recebem mais recursos, indiferentemente de outros critérios. A Figura 5, acima, endossa esse cenário, haja vista que as regiões mais pobres do país são desprestigiadas em relação às mais ricas em termos de quantidade de instrumentos celebrados.

2.2.O papel do uso da tecnologia aplicada ao Controle

Diante do exposto, com amparo na revisão bibliográfica referenciada, parte-se da premissa que as novas tecnologias de informação são um importante instrumento para o exercício da atividade de auditoria interna governamental.

Dunleavy et al (2006) argumentam que o *New Public Management* tende a morrer e dar espaço a uma nova era de governança digital, denominada “*Digital-Era Governance*”. Esse novo paradigma se vale dos avanços tecnológicos e tem como objetivo torná-los centrais nas ofertas dos serviços públicos. Ou seja, a tecnologia deixa de ser um instrumento que possibilita ganhos de eficiência para se tornar uma solução que ajuda a moldar o próprio serviço público.

Em primeira análise, o *New Public Management* deixou de ser novo (Barzelay 2019) configurando-se, reconhecidamente, em um modelo de meia-idade, o qual, aplicado na realidade burocrática contemporânea, pode gerar adversidades nos resultados esperados. Torna-se urgente reverter esse déficit, enfatizando que a Tecnologia da Informação e os Sistemas de Informação são mecanismos chaves para a continuidade da racionalização weberiana dos processos da administração pública na década de 2020. O futuro chegou. Naturalmente, destaca-se a grande importância que a tecnologia assume na nova maneira de operar os sistemas de gerenciamento e dos métodos de interagir com os cidadãos e os consumidores de serviço público (Barzelay e Gallego 2006)

As raízes do desprestígio do *New Public Management* são fundadas na negligência desse modelo em reconhecer que o empecilho para a simplificação da Administração Pública perpassa pela superação da conjuntura de “criar dificuldades para vender facilidades” (Bresser-Pereira 1996). O problema se manifesta quando servidores públicos exercendo o papel de auditores instauram burocracias-boutique em que criam complexos níveis de exigências formais. Perverter a racionalidade para desenvolver fortes gerências legalistas nas agências governamentais são vistas como problemáticas, porque geram atitudes de gerenciamento obcecadas por objetivos organizacionais intermediários, ao invés de prestação de serviços ou eficácia (Hood, James e Scott 2000; Secchi 2009).

Não obstante, o modelo “*Digital-Era Governance*” influencia a atividade de auditoria e controle. Isso se deve porque a auditoria e o controle (Dunleavy et al. 2006) são parte da ciência social Administração (Fayol 1949; Mintzberg 1979; Taylor 1911). Por seu turno, essa

influência implica em efeitos abrangentes uma vez que não está restrita a aplicações iminentemente tecnológicas (Newell e Simon 1976). Diferentemente do que se sugere, a influência do “*Digital-Era Governance*” na atividade de auditoria não se restringe aos instrumentos de tecnologia, pois essa influência provoca a associação de uma ampla arquitetura de conceitos envolvendo cognição, comportamento, mudança organizacional, política e cultural ligada a tecnologia da informação. O advento “*Digital-Era Governance*” é atualmente a influência mais difundida e peculiar sobre como os arranjos de governança estão se modificando para o exercício do controle governamental.

Percebe-se que o paradigma vigente faz com que os auditores se furtem em adotar técnicas obsoletas (Vasarhelyi e Halper 1991) como exames físicos de papel, exame de autenticidade de documentos por meio de comparação visual ou correlacionar informações por meio de observação presencial. Da mesma forma em que datilógrafos foram extintos da Administração Pública porque não se usa mais as máquinas de escrever, os auditores devem estar atentos para não terem o mesmo fim. As técnicas de auditoria ultrapassadas devem ser deixadas de lado, ao passo que a oferta de serviços públicos é preponderantemente digital, exigindo que o servidor público investido na atividade de auditoria opere sistemas de informação eletrônicos. Nesse aspecto, surge a necessidade de enfrentamento das novas relações do ecossistema de disrupção tecnológica (Margetts e Dunleavy 2002) criado pelo “*Digital-Era Governance*”.

Segundo Parker, Jacobs e Schmitz (2019) quanto mais inovadora a tecnologia, maior a perspectiva de que os auditores proponham novos paradigmas para a operação do negócio e para as tomadas de decisão do gestor. Por sua vez, quanto maior a complexidade da tecnologia, maior a possibilidade de inovação, mas a dificuldade de sua implementação é majorada, justamente pela complexidade de seu manuseio pelos auditores. Essa implementação de novas tecnologias exige mudança de abordagem e comportamento. Toda ação provoca uma reação. Consequentemente, instaura-se uma resistência na adoção de novas tecnologias por parte dos auditores (Kim, Mannino, e Nieschwietz 2009).

Naturalmente, o aparato burocrático brasileiro exerce e continuará exercendo resistência para mudar a lógica de aceitação de inovações (Heber, 2014) em que a regra é a interpretação obtusa do princípio da legalidade (Campana 2017). Afinal de contas, acata-se ao pé da letra “*Dura Lex, Sed Lex,*” a lei é dura, mas é a lei. No contexto público só é permitido fazer aquilo que está nos limites da legislação, enquanto qualquer ato que extrapole as regras normativas pode ser considerado ilegal (Bresser-Pereira 2008). Isso implica em duas vertentes nefastas:

inibe o bom auditor a propor alternativas não tradicionais e inovadoras, e faz com que o mal auditor esteja amparado com justificativas plausíveis para sua inércia.

Todavia, segundo os autores Kim, Mannino e Nieschwietz (2009) há progressos na aceitação de tecnologias na atividade de auditoria interna, fazendo com que uma abordagem inovadora ancore a gestão.

Adicionalmente, a cultura do medo, disseminada na gestão pública (Alves e Calmon 2008), foi um fator preponderante, e continua sendo, para a plenitude do desenvolvimento da prática inovadora, o sistema "malha fina de convênios". No âmbito da Administração Pública, observa-se atualmente a formação de uma contenda em que os pólos litigantes são de um lado os gestores públicos e de outro os órgãos de controle. Por seu turno, a configuração desse cenário é muito maléfica, na medida em que suas consequências são nefastas para a engrenagem da máquina pública (Power, 2003b, 2003a).

Precipuamente, o controle tem como objetivo primordial assegurar a boa governança (Brasil 2017a; COSO 2013). Contudo, gestores públicos frequentemente apontam que os órgãos de controle transpõem as atribuições a que se destinam, causando interferências inapropriadas na gestão. Por sua vez, essas interferências pautam a gestão, haja vista que causam alguns efeitos nos gestores: receio na execução de orçamento, omissão para implementar inovação, aversão para assumir riscos, e por fim inquietação em assumir cargos para executar as políticas públicas.

Diante do exposto, configura-se o paradoxo do controle: sua atuação aprimora a governança pública à medida que causa efeitos adversos na gestão. Dessa forma, qual seria a relação das ações de controle com a gestão orientada ao medo?

O triunfo sobre esse obstáculo que se postou em frente à implementação do “*malha fina de convênios*”, qual seja, o medo dos gestores em definir um apetite ao risco para estarem aptos a usar o malha de convênios e então aprovarem contas automaticamente, exigiu o uso da persistência como um fator chave (Power 2009). De fato, a proposta de um novo arquétipo procedimental para análise instantânea de prestação de contas gera desconforto nos gestores quando esses o comparam com o processo tradicional de prestação de contas.

O processo tradicional de prestação de contas envolve manipulação de processos, que muitas vezes são constituídos de papel físico, elaboração de pareceres, diligências com o conveniente, verificação de recibos e notas fiscais, tramitação na cadeia hierárquica, e está sujeito a todas as intempéries de um processo administrativo comum de uma burocracia

racional (Ramos 1983; Tenorio 1981). A alternativa de uma análise que gera um resultado instantâneo, eximindo as etapas necessárias de um processo tradicional, é vista como uma revolução pelo tomador de decisão. Porém, gera receio.

Dessa forma, a persistência e resiliência adotada pela equipe do "Malha Fina de Convênios" foram essenciais para não permitir o desvanecimento da iniciativa. Logo, os principais fatores que contribuíram para implantação do “*Malha Fina De Convênios*” foram persistência, resiliência e trabalho em equipe com o Departamento de Transferências Voluntárias da União do Ministério da Economia.

2.3.A Sofisticação dos métodos de Auditoria: Técnicas de mensuração de desempenho de *Machine Learning*

A concepção e a natureza das técnicas de auditoria estão se modificando em decorrência do surgimento do “*Digital-Era Governance*”, sustentadas por frameworks como Big Data, Internet das Coisas (IoT) e Inteligência Artificial. Nesse novo paradigma, o aprendizado de máquina (*Machine Learning*) emerge com uma técnica revestida de protagonismo na atividade de auditoria (Brown-Liburd e Vasarhelyi 2015), na medida em que é muito versátil para correlacionar causa e efeito, sofisticando a metodologia de predição de resultados. Stanistic, Radojevic e Stanic (2019) afirmam que uma linha de pesquisas precedentes se baseavam no clássico uso de modelos estatísticos, principalmente regressão logísticas e probit, para desenhar uma metodologia de predição. Por sua vez, esses mesmos autores alegam que pesquisas mais recentes demonstram que o uso de técnicas de Machine Learning, em detrimento de modelos estatísticos, possui desempenho melhor na predição.

Há diversos algoritmos diferentes de aprendizado de máquina, sendo que a principal distinção entre eles está nas categorias de aprendizado supervisionado e não-supervisionado. Entre os algoritmos supervisionados, destaca-se o *Random Forest*, uma vez que apresenta o melhor desempenho em comparação com outros algoritmos supervisionados, entre eles Naive Bayesian, Bayesian Belief Network, Artificial *Neural Network*, *Support Vector Machine*, *Decision Tree* e *Random Tree*. Oportunamente, essa constatação empírica encontra amparo em Breiman (2001), Cecchini et al. (2010), Tiwari e Hooda (2018) e Grover, Bauhoff e Friedman (2019), na medida em que esse afirma que o *Random Forest* constrói um mecanismo de árvores de decisão múltiplas e seleciona o melhor cenário de combinação de variáveis para fornecer uma predição e por isso possui um desempenho melhor entre as outras categorias de algoritmo

de aprendizado de máquina. O *Random Forest* utiliza a combinação de uma miríade de árvores de decisão para computar um *score*, que significa o grau de previsão de classificação. Esse *score* é calculado com base na média das previsões de cada árvore individual. Conseqüentemente, o *Random Forest* fornece uma nota de risco mais assertiva para medir a probabilidade de aprovação ou reprovação das contas, por exemplo.

Por seu turno, a precisão de algoritmos de classificação estão atrelados ao seu desempenho (Fawcett 2006). Na literatura, existem várias abordagens sobre a aferição de desempenho de algoritmos de aprendizado de máquina, sendo que Stanisic, Radojevic e Stanic (2019), Zhang (2019a) e Bao et al. (2020) sublinham que não há uma consagrada. Nesse panorama, uma abordagem comum para avaliar o desempenho é proceder com testes de validação cruzada (Kohavi 1995). Todavia, os dados utilizados para treinamento, quais sejam convênios reprovados e aprovados ao longo dos anos, são de natureza intertemporal, o que torna a aferição de desempenho por meio da validação cruzada inadequada (Bao et al. 2020). Oportunamente, a predição do resultado da análise das prestações de contas pode ser categorizada em duas classes mutuamente exclusivas: aprovada e rejeitada. Conseqüentemente, torna-se viável a mensuração de desempenho por meio de métricas que aferem eventuais erros de categorização nas classes (i) prestação de contas aprovada ou (ii) prestações de contas rejeitada (Bao et al. 2020; Stanisic, Radojevic e Stanic 2019; Zhangb 2019a).

Nessa configuração específica em que os algoritmos de aprendizado de máquina classificam as instâncias, neste caso convênios, em duas categorias exclusivas (aprovado ou rejeitado), Alpaydin (2020) e Fawcett (2006) propõem uma abordagem analítica para problemas de classificação quando existem apenas duas categorias. Nesse cenário há quatro resultados possíveis. De um lado, se a predição do algoritmo for correta para um convênio aprovado, temos um verdadeiro positivo; caso a predição for incorreta para convênio aprovado, há um falso negativo. Por outro lado, se a predição do algoritmo indicar corretamente um convênio reprovado, temos um verdadeiro negativo; caso a predição apontar incorretamente um convênio reprovado, há um falso positivo. Assim, o conjunto dos resultados das predições do algoritmo classificador possibilita a elaboração de uma tabela de contingência com duas linhas e duas colunas, comumente denominada matriz de confusão, que representa as disposições do conjunto de resultados. A matriz de confusão permite identificar quais tipos de classificação incorreta ocorrem, ou seja, quais as categorias (convênios aprovados ou reprovados) são frequentemente confundidas.

Figura 6 – Matriz de Confusão.

	<i>Resultado Aprovado</i>	<i>Resultado Rejeitado</i>
<i>Hipótese de Aprovação</i>	Verdadeiro Positivo (VP)	Falso Positivo (FP)
<i>Hipótese de Rejeição</i>	Falso Negativo (FN)	Verdadeiro Negativo (VN)

Adaptado de Fawcett (2006) e Alpayadin (2020).

A Figura 6 mostra uma matriz de confusão. Os números na diagonal principal (VP e VN) representam as previsões corretas do algoritmo, e os números da diagonal secundária (FP e FN) representam os erros, a confusão entre as duas categorias. Por sua vez, a matriz de confusão forma a base de 3 métricas para aferição de desempenho de algoritmos de classificação gerados por *machine learning*: **acurácia, especificidade e sensibilidade**. Essas métricas são amplamente utilizadas na literatura, conforme se observa Alpayadin (2020), Fawcett (2006), Hooda, Bawa, e Rana (2020), Grover, Bauhoff e Friedman (2019), Tiwari e Hooda (2018) e Zhang(2019a), não obstante Hooda, Bawa, e Rana (2020) e Tiwari e Hooda (2018) adotarem uma forma de computa dessas métricas diferente dos demais. A metodologia de cálculo da acurácia, especificidade e sensibilidade adotada majoritariamente na literatura é:

$$\text{Acurácia} = \frac{VP + VN}{VP + VN + FN + FP}$$

$$\text{Sensibilidade(Recall)} = \frac{VP}{VP + FN}$$

$$\text{Especificidade} = \frac{VN}{VN + FP}$$

Contudo, as métricas para mensuração de desempenho usualmente utilizadas (acurácia,

especificidade e sensibilidade) não são apropriadas quando há um balanceamento desequilibrado da sequência de treinamento, em que a ocorrência de um categoria é muito maior que a outra (Grover, Bauhoff e Friedman 2019; Bao et al. 2020; Alpaydin 2020).

Segundo Bao et al. (2020), poder-se-ia utilizar uma métrica, denominada precisão balanceada, como uma alternativa para aferir adequadamente o desempenho de um modelo preditivo. Essa métrica balanceada é definida como a média entre a média simples entre a sensibilidade e a especificidade de predição das instâncias (prestações de contas de convênios). Por seu turno, a métrica de precisão balanceada é:

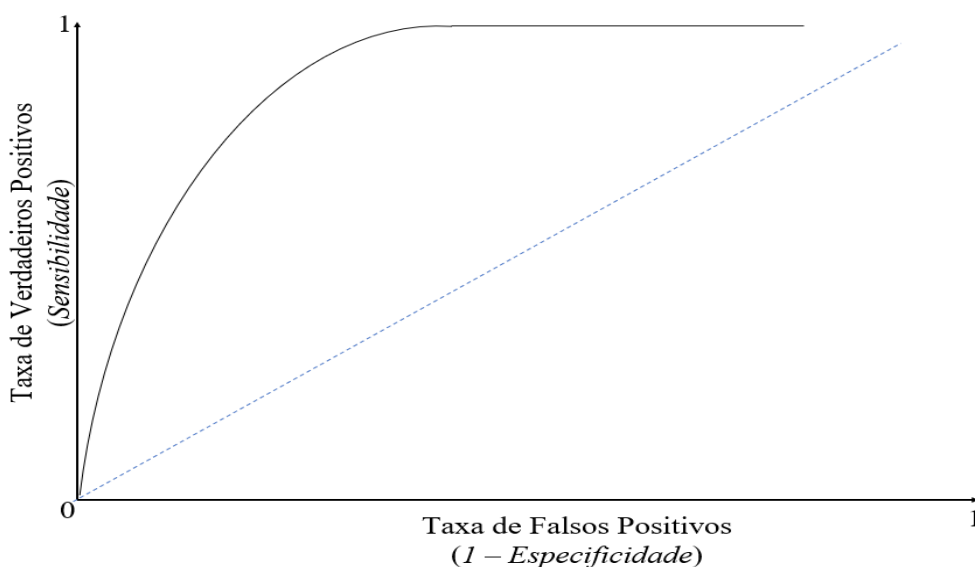
$$\text{Precisão Balanceada} = \frac{\text{Sensibilidade} + \text{Especificidade}}{2}$$

Por outro lado, Larcker e Zakolyukina (2012) ponderam que há uma limitação no uso da métrica de precisão balanceada como parâmetro de avaliação de desempenho. A referida métrica é concebida com base na assertividade de predição das classes, no caso em tela, aprovação ou rejeição. Objetivamente, isso significa que a predição depende da rotulação dos dados na sequência de treinamento, ou seja, o resultado conclusivo da análise que os órgãos concedentes realizaram nas prestações de contas apresentadas. Ocorre que os algoritmos de aprendizagem são bastante sensíveis ao desequilíbrio da sequência de treinamento. Uma vez que os órgãos concedentes reprovam um número muito reduzido de convênios proporcionalmente ao total, isso influenciará na distribuição da predição, conseqüentemente na sensibilidade e especificidade, superestimando-as (Fawcett 2006).

Destarte, uma maneira de superar essa limitação é a utilização da curva ROC (*Receiver Operating Characteristics*) e aferição da métrica AUC (*Area Under the Curve*) (Bao et al., 2020; Fawcett, 2006; Stanisc, Radojevic e Stanic 2019). A curva ROC é uma representação bidimensional do desempenho de um classificador, nesta pesquisa representado pelo algoritmo do “*Malha Fina de Convênios*”. Por seu turno, a curva ROC combina a taxa de ocorrência de verdadeiros positivos, a sensibilidade, e a taxa de ocorrência de falsos positivos, o valor da subtração de 1 e a especificidade, em um gráfico com dois eixos. Por conseguinte, é possível inferir o desempenho de um modelo preditivo por meio de um único escalar ao calcular a métrica AUC. Como a AUC é uma parte da área do quadrado da unidade, seu valor sempre estará entre 0 e 1,0. Na medida em que uma seleção aleatória produz a linha diagonal entre (0,

0) e (1, 1), que tem uma área de 0,5, nenhum algoritmo de predição plausível deveria ter uma AUC menor que 0,5. O valor até 0,5 significa um modelo sem qualquer poder discriminativo e o valor 1 significa um modelo perfeito. A métrica AUC equivale à probabilidade de que um convênio reprovável escolhido aleatoriamente seja classificado com um score mais alto (verdadeiro negativo) do que um convênio aprovável (verdadeiro positivo), também escolhido aleatoriamente (Bao et al. 2020; Breiman 2001; Fawcett 2006; Stanistic, Radojevic e Stanic 2019;).

Figura 7 – Curva ROC.



Adaptado de Alpaydin (2020) e Fawcett (2006). Elaborado pelo Autor.

As curvas ROC são uma ferramenta muito útil para visualizar o desempenho de algoritmos de classificação (Cecchini et al. 2010). Elas são capazes de fornecer percuciência da medida do desempenho, complementarmente às medidas escalares como acurácia, sensibilidade, sensibilidade ou taxas de erro. No entanto, Fawcett (2006) sublinha que a curva ROC deve ser usada com parcimônia, pois, como com qualquer métrica de avaliação, possui limitações. Dessa forma, é sugerido que sua adoção seja cotejada com as outras métricas escalares.

3. CONTEXTUALIZAÇÃO DO PROBLEMA

O objetivo desta seção é tratar do gargalo operacional que se configurou no fluxo processual das Transferências Voluntárias da União. A criticidade do problema da falta de capacidade na etapa de análise de prestação de contas é elucidada, com o destaque de que esse problema se agrava com a contínua ampliação do quantitativo de prestações de contas pendentes de análise conclusiva. Esta seção está dividida em mais 4 partes, para além desta introdução. Na primeira descreveremos os incentivos que levaram à concepção do “*Malha Fina de Convênios*”. Na segunda parte, tratamos do que consiste essa solução. Em seguida, na terceira parte, são colocados os problemas e as incertezas na sua utilização. Por fim, na quarta seção, evidenciamos a importância desses problemas, a ponto de motivarem este trabalho.

3.1. Transferências Voluntárias e a Plataforma +Brasil (antigo SICONV)

O processo de Transferências Voluntárias da União movimentou mais de R\$ 135 bilhões entre 2008 e 2019 por meio de mais 170 mil instrumentos entre os entes da Federação, segundo dados da Plataforma +Brasil. Precisamente, em novembro de 2019 contava com 175.456 instrumentos cadastrados (Brasil, 2019b).

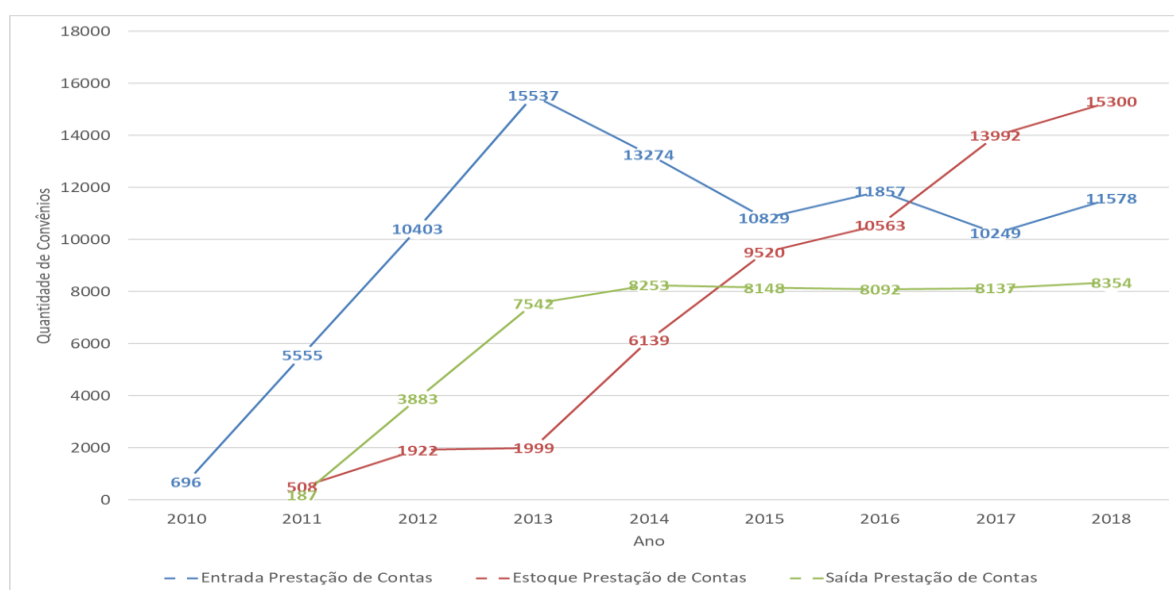
Contudo, a concessão de transferências voluntárias pelos órgãos e entidades governamentais constitui um grande desafio aos administradores públicos, tanto no tocante à lisura desejada, quanto à operacionalização ágil e eficaz dos milhares de instrumentos destinados à implementação de políticas públicas nos Estados e Municípios brasileiros. Esse desafio ocorre devido a quantidade de agentes e a multiplicidade de objetos presentes na pactuação da concessão.

Entre as políticas públicas que receberam recursos por meio de repasse federal nos últimos 10 anos elenca-se: planejamento, gestão e desenvolvimento urbano (R\$ 11,3 bi); apoio ao turismo (R\$ 9,3 bi); atenção especializada em saúde (R\$ 7,5 bi); agricultura familiar (R\$ 5,5 bi); recursos hídricos (R\$ 4,3 bi); segurança pública (R\$ 4,3 bi); políticas indígenas (R\$ 4,2 bi); esportes de grandes eventos (R\$ R\$ 3,7 bi); esporte e lazer (R\$ 2,8 bi); e agricultura e pecuária (R\$ 2,7 bi).

Em que pese as deficiências notadamente reconhecidas na Plataforma +Brasil (antigo SICONV), esse sistema vem sendo aprimorado ao longo da última década e hoje permite

atender requisitos primordiais da boa governança pública, a qual exige a transparência como pilar fundamental. Por sua vez, o processo de transferências voluntárias se mostra ainda excessivamente moroso, pouco eficiente e eficaz (Brasil 2018a). Segundo os dados de um relatório de auditoria da CGU, o tempo médio do ciclo de vida de um processo de repasse de verbas federais no SICONV atinge mais de 5 anos. Adicionalmente, a CGU detectou um grande desequilíbrio entre a capacidade operacional dos órgãos concedentes e o volume de trabalho despendido para analisar as prestações de contas dos repasses efetuados. A quantidade de repasses efetuados exigiria um esforço de análise muito superior à capacidade de análise disponível dos órgãos repassadores.

Figura 8 - Estoque de prestação de contas pendentes de análise.



Elaborado pelo Autor. Adaptado de Brasil (2018a)⁵. Fonte: Plataforma +Brasil.

Diante desse gargalo operacional, emerge o problema de estoque de prestação de contas pendentes de análise pelos órgãos repassadores. A criticidade do problema é caracterizada pelo crescimento contínuo desse estoque. Na perspectiva da atividade de auditoria governamental,

⁵ A Figura 8 foi elaborada a partir do histórico da base(fotografia) no último dia de cada ano. Apenas assim seria possível verificar a real situação ano-a-ano da situação dos convênios, na medida em que a base de dados da Plataforma +Brasil possui inconsistências. Ao longo de 10 anos, muitas intervenções e atualizações no sistema Plataforma +Brasil se sucederam. Dessa forma, não seria fidedigno ilustrar a evolução do passivo de prestação de contas pendente de análise por meio de uma simples conta aritmética, subtraindo a quantidade de prestação de contas submetidas para análise em um determinado ano pela quantidade que foram analisadas no ano subsequente.

há duas linhas de atuação (Brasil 2017; IIA 2012; Olivieri 2011) para a situação ilustrada na Figura 8.

A primeira linha é a avaliação independente e objetiva do processo de operação das transferências voluntárias visando o aprimoramento do gerenciamento de risco, dos controles internos e da governança (Brasil 2017a). O resultado imediato de atuação nessa abordagem seriam algumas recomendações cujo teor demandaria que os órgãos repassadores de recursos envidem esforços para realizarem a análise de prestação de contas, ou caso não tenham capacidade de análise, abstenham-se de promover novos repasses. Trata-se de recomendações relevantes e úteis, porém óbvias, não passa de mais do mesmo (Power 2003b).

Por outro lado, a segunda linha de atuação consiste na consultoria, em que a auditoria presta serviços de assessoria e aconselhamento em assuntos estratégicos (Brasil 2017a) do processo de transferências voluntárias da união. O resultado da consultoria, geralmente, é um produto ou uma solução construída conjuntamente com o gestor. No caso em epígrafe, a consultoria produziu uma solução para o estoque de prestação de contas de repasses pendentes de análise.

Após a concretização da realização da auditoria interna seguindo os preceitos da forma de atuação em consultoria, foi entregue um sistema que permite resolver o flagrante desequilíbrio entre a capacidade operacional dos órgãos concedentes e o volume de trabalho requerido para analisar a prestação de contas. Esse desequilíbrio gerou um passivo de mais de 15 mil instrumentos, os quais correspondem a aproximadamente R\$ 17 bilhões pendentes de análise. Esse sistema foi denominado malha fina de convênios e consiste no uso de inteligência artificial.

Com base nas características de cada convênio ou contrato de repasse, a ferramenta reconhece padrões e permite prever, com elevado grau de precisão, o resultado da análise de contas, caso fosse realizada uma avaliação manual por servidores dos órgãos federais concedentes (Huang e Vasarhelyi 2019; Issa, Sun, e Vasarhelyi 2016). Na prática, a aplicação da “Malha Fina de Convênios” verifica os instrumentos operacionalizados na Plataforma +Brasil, utiliza o aprendizado de máquina com o algoritmo supervisionado *Random Forest* e fornece um *score* para medir a probabilidade de aprovação ou reprovação das contas. A metodologia combina também a emissão de alertas gerados nas trilhas de auditoria aplicadas pela CGU, na busca por padrões pré-definidos de indícios de impropriedades ou irregularidades.

3.2. O Sistema “*Malha Fina de Convênios*”

O sistema “*Malha Fina de Convênios*” é um modelo preditivo criado pela CGU que permite indicar, com determinado grau de certeza, o resultado da análise das prestações de contas dos convênios quando são apresentadas pelos convenentes (entes subnacionais recebedores de recursos) aos concedentes (órgãos da União repassadores de recursos). Ou seja, o Malha Fina permite inferir se as contas dos convênios serão aprovadas ou rejeitadas.

Resumidamente, esse sistema foi o resultado do desenvolvimento de uma solução que usa um algoritmo de aprendizado de máquina com base nas características dos convênios cujas contas já foram analisadas. Entre setembro de 2008 e dezembro de 2017 mais de 61.000 convênios (Brasil 2019b) tiveram suas contas analisadas pelos concedentes, proporcionando uma quantidade de dados satisfatória para que o aprendizado do algoritmo fornecesse resultados com precisão. Ato contínuo, 104 variáveis de cada convênio foram utilizadas no algoritmo de aprendizagem supervisionado *Random Forest*, por seu desempenho ser melhor em comparação com outros (Breiman 2001; Domingos 2012).

Igualmente, o “*Malha Fina de Convênios*” possibilitará maior agilidade na deflagração do gatilho das providências administrativas a serem tomadas visando à elisão do dano ao Erário. Entre 2018 e 2019, foram instaurados 2.869 processos de Tomadas de Contas Especial (TCE) relativos a convênios e contratos de repasse com indícios de irregularidades, que buscam recuperar prejuízos estimados em mais de R\$ 6 bilhões⁶.

Esse sistema foi oficialmente disponibilizado à Administração Pública Federal em 07/11/2018 com a publicação no Diário Oficial da União da Instrução Normativa Interministerial nº 5, de 6 de novembro de 2018 (Brasil 2018b), editada conjuntamente entre a CGU, o extinto Ministério da Fazenda e o extinto Ministério do Planejamento, Desenvolvimento e Gestão. Aludida Instrução Normativa estipulou a data de 31/08/2018 como delimitador de prazo para definir o passivo de estoque. Dessa forma, todas as prestações de contas enviadas aos órgãos concedentes até 31/08/2018 são considerados estoque e sua análise pode ser amparada pelo sistema malha fina de convênios (Tabela 2).

Em seguida, a Instrução Normativa ME/CGU nº 1, de 14 de fevereiro de 2019 (Brasil, 2019a), ampliou a abrangência de aplicabilidade do “*Malha Fina de Convênios*” para

⁶ Dados extraídos do sistema e-TCE, mantido pela Controladoria-Geral da União e Tribunal de Contas da União.

prestações de contas submetidas para análise a partir de 01/09/2018, conquanto o estoque de passivo referente à data base de 31/08/2018 esteja saneado.

Destarte, as duas Instruções Normativas supracitadas delinearão o escopo de abrangência do “*Malha Fina de Convênios*” e estabelecerão o marco de estreia. A operação do sistema na Plataforma +Brasil deu início a partir da publicação da IN MP/MF/CGU nº 05, de 06/11/2018. Essa Instrução Normativa contribuiu de forma significativa para a redução do prazo médio da fase de prestação de contas, que em 2018 superava em 2,5 anos. Com a inovação, houve um benefício imediato decorrente da diminuição de custos de processos administrativos, como a quantidade de servidores relacionados à análise do passivo existente (Issa, Sun, e Vasarhelyi 2016; Wirtz, Weyerer, e Geyer 2019).

Órgão Superior Concedente	Total de Convênios		Valor Total dos Convênios	
Ministério da Saúde	2243	14,66%	R\$ 2.064.572.925,18	12,36%
Ministério do Turismo	2115	13,82%	R\$ 1.070.053.176,45	6,41%
Ministério do Esporte	1416	9,25%	R\$ 1.301.841.452,22	7,80%
Ministério da Justiça	1205	7,88%	R\$ 1.302.123.217,50	7,80%
Ministério da Integração nacional	1109	7,25%	R\$ 1.023.379.009,39	6,13%
Ministério da Educação	1059	6,92%	R\$ 1.444.256.052,34	8,65%
Ministério da Agricultura, Pecuária e Abastecimento	894	5,84%	R\$ 516.133.850,15	3,09%
Presidência da República	830	5,42%	R\$ 495.708.688,49	2,97%
Ministério da Cultura	823	5,38%	R\$ 587.646.518,43	3,52%
Ministério das Cidades	760	4,97%	R\$ 371.683.423,51	2,23%
Ministério do Desenvolvimento Social	653	4,27%	R\$ 2.879.625.133,53	17,24%
Ministério dos Direitos Humanos	545	3,56%	R\$ 370.170.155,79	2,22%
Secretaria de Agricultura Familiar e do Desenvolvimento Agrário	469	3,07%	R\$ 648.547.655,40	3,88%
Ministério do Desenvolvimento Agrário	305	1,99%	R\$ 473.913.885,79	2,84%
Ministério do Trabalho e Emprego	258	1,69%	R\$ 518.660.571,00	3,11%
Ministério da Defesa	230	1,50%	R\$ 637.759.438,63	3,82%
Ministério da Ciência, Tecnologia, Inovações e Comunicações	192	1,25%	R\$ 649.926.740,64	3,89%
Ministério do trabalho e Previdência Social	70	0,46%	R\$ 138.538.169,93	0,83%
Ministério do Meio Ambiente	53	0,35%	R\$ 39.800.858,50	0,24%
Ministério do Desenvolvimento, Indústria e Comércio Exterior	52	0,34%	R\$ 146.587.135,28	0,88%
Ministério de Minas e Energia	16	0,10%	R\$ 15.770.059,31	0,09%
Ministério dos Transportes	1	0,01%	R\$ 1.430.704,88	0,01%
Justiça Eleitoral	1	0,01%	R\$ 677.247,82	0,00%
Ministério do Planejamento, Desenvolvimento e Gestão	1	0,01%	R\$ 777.825,00	0,00%
Total Geral	15300	100,00%	R\$ 16.699.583.895,16	100,00%

Tabela 2 - Passivo de convênios em 31/08/2018. Elaborado pelo Autor. Fonte Plataforma +Brasil..

A necessidade de apresentação de uma solução para a adversidade imposta pelo crescente aumento de convênios sem prestação de contas conclusiva foi encorajada pelo pleito recorrente dos gestores envolvidos na operação dos convênios. Mesmo antes da instituição do antigo SICONV em 2007, o problema do passivo de prestação de contas em convênio tem sido um percalço inconveniente aos órgãos concedentes e aos órgãos de controle que rotineiramente precisam flexibilizar as exigências impostas de prazos para conclusão de prestação de contas. Dessa forma surge a encomenda para a realização de pesquisas e critérios que pudessem amparar a análise de prestações de contas de modo mais célere (Brasil 2017a, 2017b) tendo em vista o constante incremento da quantidade de instrumentos em estoque.

O principal objetivo da “*Malha Fina de Convênios*” é solucionar o problema crítico de falta de capacidade operacional que envolve o processo de transferências voluntárias da União. A situação vem se agravando, ao considerar a contínua ampliação do quantitativo de prestações de contas pendentes de análise. Até o final do mês de agosto de 2018, a quantidade de convênios inseridos em um estoque de passivo aguardando prestação de contas já somava 15,3 mil instrumentos, no valor total de R\$ 16,7 bilhões.

O público alvo desta iniciativa são todos os órgãos concedentes de recursos de transferências discricionárias operacionalizadas na Plataforma +Brasil por meio de convênios e contratos de repasse. Especificamente, os setores que mais se beneficiaram ou poderão se beneficiar do sistema malha fina de convênios são aqueles responsáveis pela execução financeira e análise de prestação de contas dentro dos órgãos. Geralmente, a estrutura administrativa que vigora no aparato burocrático do governo federal denomina esses setores como Diretoria de Gestão Interna, Diretoria de Finanças e Contabilidade, Subsecretaria de Planejamento, Orçamento e Administração, Diretoria de Administração, entre outros. Assim, o público alvo determinante do “*Malha Fina de Convênios*” são 140 órgãos da Administração Pública Federal Direta, Autárquica e Fundacional que celebram convênios sobre a égide das normas da Plataforma +Brasil.

A essência do “*Malha Fina de Convênios*” reside na aplicação de um algoritmo que atribui um *score* individual para cada convênio, variando entre 0 e 1. Quanto mais próximo de 0 estiver o *score*, maior a chance de o convênio ter suas contas reprovadas. Alternativamente, quanto mais próximo de 1, maior as chances de o convênio ter suas contas rejeitadas. Conseqüentemente, a rejeição das contas de um convênio enseja o concedente a tomar as medidas cabíveis para a recuperação do dano ao Erário, como por exemplo, a instauração de

uma Tomada de Contas Especial (TCE).

Adicionalmente, cabe registrar que o sistema “Malha Fina de Convênios” vai ao encontro das premissas estabelecidas no modelo “*Digital-Era Governance*” e endossa a ruptura em relação ao “*New Public Management*” (Dunleavy et al. 2006). Há um ganho de eficiência ao implementar o “Malha Fina de Convênios”, ao passo que o trabalho rotineiro dos servidores públicos que analisam convênios passa a ser racionalizado weberianamente, em seu sentido estrito, já que a partir dos resultados oriundos do processamento de dados realizado pelo sistema, serão determinados quantos e quais convênios devem ser avaliados mais detidamente pelas pessoas.

Logo, o “*Malha Fina de Convênios*” não se reveste apenas de uma inovação tecnológica que oportuniza, por exemplo, a digitalização das prestações de contas, e que torna o trabalho mais eficiente por agilizar a prestação de contas por meio do tratamento e a análise dos dados (outrora visão do *New Public Management*); mas sim, de um novo papel dado à tecnologia que faz parte da política pública em si, neste caso das Transferências Voluntárias da União (Margetts e Dunleavy 2002; Barzelay 2018; Bresser-Pereira 1998).

3.3. Descrição do problema

O “*Malha Fina de Convênios*” permitiria auxiliar o trabalho de análise das prestações de contas, uma vez que esse sistema reveste-se de todas as qualidades do que Huang e Vasarhelyi (2019), Sun (2019) e Zhang (2019b) denominam *Robotic Process Automation* (RPA). Um RPA serve para automatizar tarefas bem-definidas e repetitivas, neste caso a análise de prestação de contas de convênios pelos órgãos concedentes. Basta que o órgão estipule uma nota mínima ante a qual todos os convênios classificados abaixo dela sejam aprovados. Contudo, existem hipóteses e condicionantes em que a adoção do “*Malha Fina de Convênios*” na análise de prestação de contas seja mais assertiva?

Uma objeção contundente contra a lógica da sistemática do “*Malha Fina de Convênios*” desponta: em que medida a precisão do algoritmo de inteligência artificial (Sun 2019; Zhang 2019) inviabilizaria sua adoção para a análise de contas automática, sem a necessidade de avaliação convencional por um servidor público do órgão concedente?

Primeiramente, para responder a essa pergunta é imprescindível medir a precisão do algoritmo conforme a variação do intervalo das notas (*scores*), as quais podem excursionar entre

0 e 1. Invariavelmente, a adoção do Malha Fina para análise das prestações de contas incidirá sobre a definição de um limiar de *score* em que se admitiria convênios aprovados. Todos os convênios com *score* acima desse limiar seriam considerados reprováveis, carecendo uma análise convencional. Dessa forma, torna-se muito relevante a distribuição das notas por seu intervalo acumulado, uma vez que a taxa de erro do algoritmo deve ser considerada em todo o intervalo a partir de um determinado limiar.

Como exemplo, observando os dados exibidos na Tabela 3 e a ilustração da Figura 9, se um determinado órgão estipular a nota 0,8 como seu limiar, significa que 79,4% dos seus convênios poderão ser passíveis de aprovação tácita, sendo que entre esses, 4,62% seriam aprovados inadvertidamente, na medida em que o algoritmo de classificação não é perfeito. Notadamente, o estabelecimento de limiar de nota por parte de órgão concedente reflete o seu apetite ao risco. Conforme as curvas da Figura 9, quanto maior o apetite ao risco, maior será nota de limiar e maior serão as chances de aprovar tacitamente convênios reprováveis.

Tabela 3 - Distribuição do intervalo acumulado de notas atribuídas pelo "Malha Fina de Convênios" dos convênios analisados entre 01/01/2018 e 18/04/2018.

Situação da Prestação de Contas	Intervalo Acumulado de <i>Score</i>						
	[0,0 a 0,4)	[0,0 a 0,5)	[0,0 a 0,6)	[0,0 a 0,7)	[0,0 a 0,8)	[0,0 a 0,9)	[0,0 a 1,0]
Aprovada	201	436	639	842	1.012	1.192	1.305
Aprovada com Ressalvas	9	21	44	84	162	264	336
Rejeitada		1	6	12	22	90	307
Total	210	458	689	938	1.196	1.546	1.948

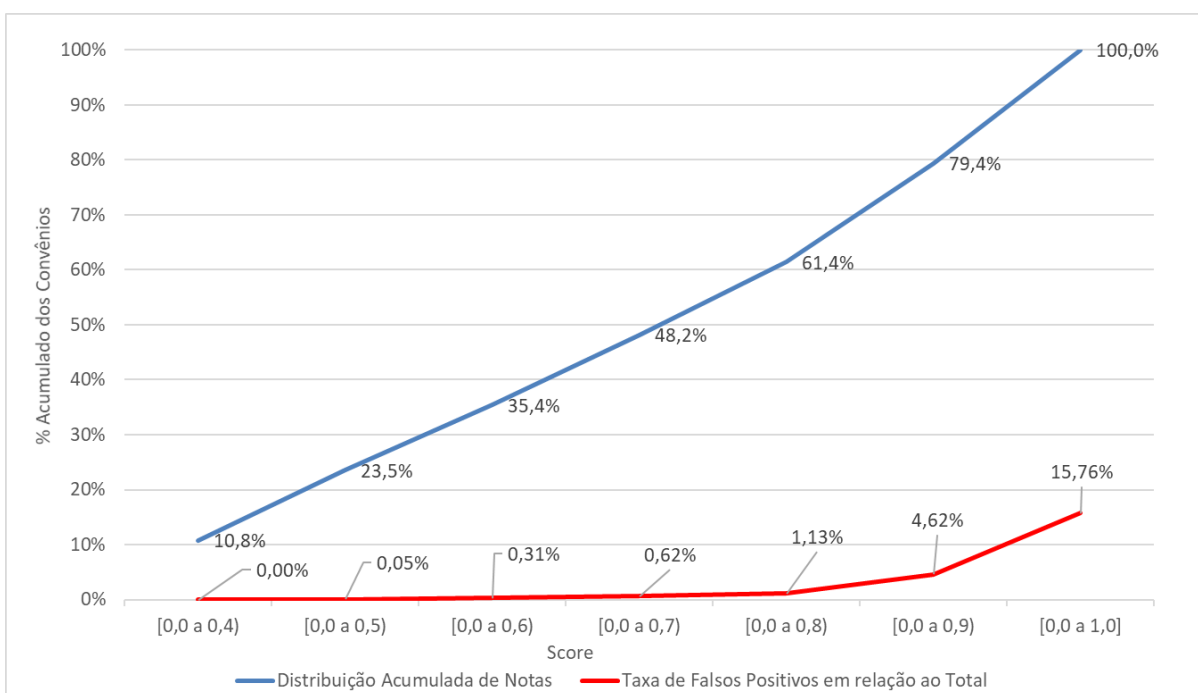


Figura 9 – Precisão do "Malha Fina de Convênios". Convênios analisados entre 01/01/2018 e 18/04/2018.

Elaborado pelo Autor.

Porém, a natureza de atuação de políticas públicas dos órgãos repassadores de recursos é plural e muito distinta entre si. As curvas apresentadas na Figura 9, contemplam todo o espectro de instrumentos celebrados na Plataforma +Brasil por órgãos da esfera federal, de maneira agregada. Ou seja, as taxas de erro exibidas não levam em consideração as qualidades intrínsecas do órgão repassador, da política pública que lhe é competente, muito menos o contexto em que se insere. Conseqüentemente, essas variáveis suscitam a possibilidade de as taxas de erro serem diferentes conforme a identidade do órgãos concedente, o que motiva a proposição de 2 hipóteses:

Hipótese 1: a precisão do algoritmo seria melhor caso o aprendizado de máquina fosse aplicado para cada órgão concedente, separadamente?

Hipótese 2: o desempenho do algoritmo apresentaria comportamento diferente caso houvesse segregação por órgão concedente?

Por sua vez, os dados obtidos sobre as taxas de reprovação de contas de convênios apresentadas pelos convenientes, restringiu-se a um período de menos de 4 meses, repercutindo entre 01/01/2018 e 18/04/2018. O algoritmo aprendeu a classificar os convênios com base em todos aqueles com contas analisadas (aprovadas, aprovadas com ressalva ou rejeitadas) desde o início de registro de dados na base da Plataforma +Brasil entre setembro de 2008 até o dia 31/12/2017.

Após a constatação que os algoritmos baseados em aprendizado de máquina por *Random Forest* erram, ao encontro do que é argumentado por Alpaydin (2020), Issa, Sun e Vasarhelyi (2016) e Sun (2019), uma questão central se estabelece sobre a volumetria da sequência de treinamento usada para o algoritmo aprender. Trata-se do ponto ótimo do tamanho da sequência de treinamento que conferiria segurança na adoção do sistema, minimizando o risco de aprovação de contas de convênios reprováveis (falsos positivos) e maximizando a análise automática, sem a necessidade de avaliação convencional por um servidor público do órgão concedente. Conforme observado na Figura 9, nos casos de avaliação realizada de maneira automática pelo sistema há uma taxa inerente de probabilidade de “erros de diagnóstico”, ou seja, convênios reprováveis, mas classificados com nota abaixo do limiar estipulado para aprovação tácita.

Nesse sentido, há a possibilidade de realizar simulações em que seria possível aferir e medir o desempenho de fidedignidade da aprendizagem conforme a agregação de dados ao longo do tempo. Por exemplo, o algoritmo poderia ser executado em lotes anuais considerando o estado da base na data de 31/12. Posteriormente, verificar-se-ia o resultado da análise das prestações de contas no ano subsequente e a comparação com as notas atribuídas pelo sistema. Logo, a perspectiva de esclarecer essas incertezas, justifica a proposição de mais duas hipóteses:

Hipótese 3: a precisão de classificação do algoritmo de aprendizado de máquina aumenta à medida que há mais dados no universo de aprendizagem?

Hipótese 4: é possível prever a existência de ponto de saturação em que a quantidade de dados não seja mais relevante para que o algoritmo de aprendizado de máquina classifique convênios?

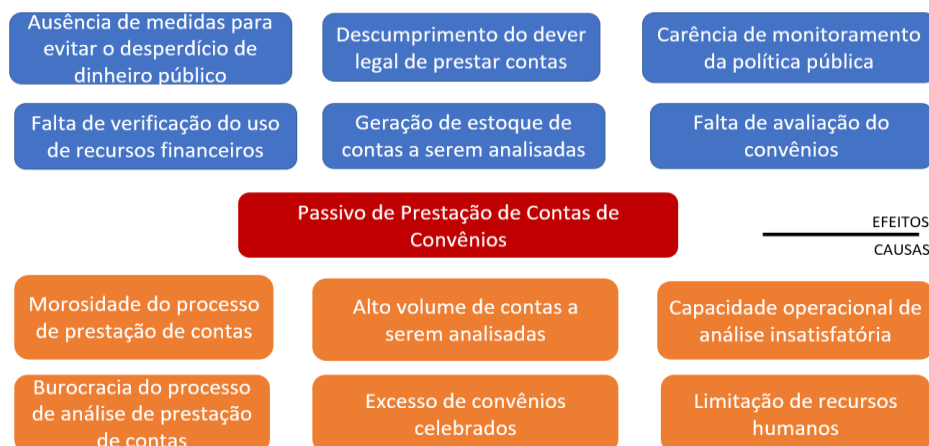
Tabela 4 - Relacionamento entre Constructos e Hipóteses de Pesquisa.

Constructo Associado	Hipóteses
<i>Identidade do repassador de recurso</i>	<p>H1: a precisão do algoritmo seria melhor caso o aprendizado de máquina fosse aplicado para cada órgão concedente separadamente?</p> <p>H2: o desempenho do algoritmo apresentaria comportamento diferente caso houvesse segregação por órgão concedente?</p>
<i>Tamanho de treinamento de dados</i>	<p>H3: a precisão de classificação do algoritmo de aprendizado de máquina aumenta à medida que há mais dados no universo de aprendizagem?</p> <p>H4: é possível prever a existência de ponto de saturação em que a quantidade de dados não seja mais relevante para que o algoritmo de aprendizado de máquina classifique convênios?</p>

3.4. Importância do Problema

O contínuo incremento da quantidade de prestações de contas pendentes de análise conclusiva possui impactos negativos gravíssimos. A existência de uma grande quantidade de convênios aguardando análise de suas prestações de contas gera consequências como: a criação de um estoque; inépcia na avaliação e no monitoramento de recursos públicos repassados aos entes subnacionais; incapacidade de a União aferir a efetividade da política pública de transferências voluntárias; e descumprimento do dever legal de prestar contas, o que reflete em total ausência de *accountability* (Schedler 1999). Com efeito, o “*Malha Fina De Convênios*” ataca frontalmente a causa dos problemas (Figura 10), o que proporcionará uma reversão dos efeitos deletérios que o estoque de prestação de contas gera na política pública das transferências voluntárias da união.

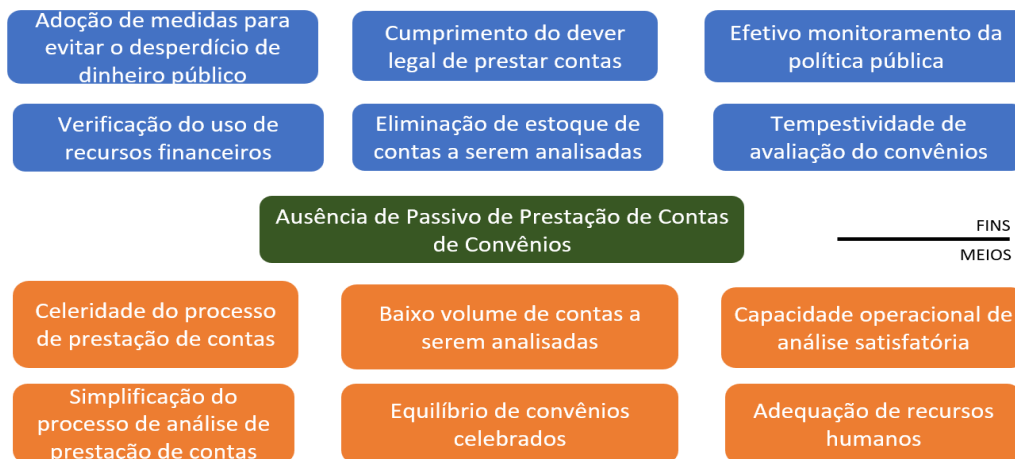
Figura 10 – Árvore de problemas da fase de prestação de contas.



Admitindo-se que o sistema “*Malha Fina de Convênios*” pode ser visto como um projeto, é de se esperar que ofereceu um resultado visando solucionar o problema do passivo mediante o emprego de um esforço temporário. Dessa forma, os meios e os recursos empregados nesse esforço temporário foram relacionados aos problemas ilustrados na Figura 10. A metodologia da árvore de objetivos aplica-se adequadamente no projeto de desenvolvimento do Malha Fina. Jackson (1997) afirma que uma vez definidos os objetivos, a análise da estratégia de ação do projeto pode se tornar mais lógica, evidenciando a relação entre os problemas elencados e os objetivos propostos. Consequentemente, as mudanças provocadas com o propósito de solucionar o problema em tela visaram a consecução de objetivos substanciais tais como: a eliminação do estoque; aptidão na avaliação e no monitoramento de recursos públicos repassados aos entes subnacionais; capacidade de a União aferir a efetividade da política pública de transferências voluntárias tempestivamente; e, por último, mas não menos importante, a observância de *accountability* por meio do cumprimento do dever legal de prestar contas (Figura 11).

Logo, a validação do método de prestação de contas automatizada é fundamental para a continuidade da abordagem inovadora proposta pelo “Malha Fina de Convênios”. Ao esclarecermos as condicionantes em que a adoção do Malha Fina na análise de prestação de contas seja mais assertiva, notadamente por meio dos constructos *Identidade do repassador de recurso* (H1 e H2) e *Tamanho de treinamento de dados* (H3 e H4), possíveis objeções contra sua implantação serão descartadas.

Figura 11 - Árvore de objetivos para solucionar o gargalo da prestação de contas.



Ademais, os benefícios advindos do uso do Malha Fina são muito relevantes para minimizar o risco da atividade de auditoria, fato que robustece ainda mais a importância do teste dos constructos propostos. Com efeito, a atividade de auditoria, como qualquer outra operação de uma organização, está suscetível ao risco de insucesso (Brasil, 2017a). Ademais, os auditores precisam tomar decisões visando a segurança na emissão de opinião daquilo que está sendo avaliado (Huang e Vasarhelyi 2019; Stanistic, Radojevic, e Stanic 2019). Todavia, essas decisões são contrapostas entre a escassez de insumos e a pressão por assegurar e certificar operações cujo universo auditável consumiria escopo, tempo e recursos muito maiores que os disponíveis (GAO 2011; Zhang 2019b). Essa dicotomia gera o risco de auditoria.

O risco de auditoria é a incapacidade do auditor em detectar um erro em uma amostra (Al-qudah, Baniahmad, e Al-fawaerah 2013). A não detecção de um erro de amostra pode ser causado por fadiga ou desatenção do auditor, como também pode ser causada pela aplicação de um procedimento de auditoria inadequado (IIA, 2012). À luz desse cenário, os mecanismos de automação robotizada e inteligência artificial que usufruem o reconhecimento de padrões para definição de eventos recorrentes configura-se em uma solução para equacionar o problema entre os escassos recursos de auditoria versus a segurança na emissão de opinião sobre a operação auditada (Power 2009). Nesta perspectiva, o “*Malha Fina de Convênios*” consiste em uma mecanismo de auditoria dessa categoria.

4. METODOLOGIA

No presente capítulo será abordada a estratégia metodológica, inspirada em Bao et al. (2020), Cecchini et al. (2010), Grover, Bauhoff e Friedman (2019), Hooda, Bawa e Rana (2020), Stanisc, Radojevic e Stanic (2019) e Zhang(2019a). Por sua vez, essa estratégia valeu-se preponderantemente da implementação empírica de um ambiente computacional em linguagem *python*. Por conseguinte, avaliar-se-á o desempenho do algoritmo do “*Malha Fina de Convênios*” (*Random Forest*) diante de diversos cenários concebidos exclusivamente para testar os constructos *Identidade do repassador de recurso* (H1 e H2) e *Tamanho de treinamento de dados* (H3 e H4).

Naturalmente, o alcance dos objetivos desta pesquisa foi viabilizado por meio das seguintes atividades:

- ✓ Mapeamento do ciclo de vida de um convênio operacionalizado no Plataforma +Brasil e relacionar todas as situações que aguardam a análise do órgão federal concedente na prestação das contas apresentada pelo ente subnacional conveniente (BRASIL, 2019b), bem como descrição do universo de convênios na Plataforma +Brasil, incluindo o corte temporal.
- ✓ Configuração de um ambiente de testes, cujo desempenho computacional fosse satisfatório, com bases de dados de diferentes datas para que o algoritmo do sistema “*Malha Fina De Convênios*” treine com diferentes séries de dados, conforme sublinha Stanisc, Radojevic e Stanic (2019) e Zhang (2019a).
- ✓ Identificação dos requisitos necessários à construção de um modelo de análise que demonstre a relação dos diferentes cenários a serem testados nas hipóteses, garantido que os recursos computacionais disponíveis não sejam empecilho para o desempenho (Breiman 2001; Domingos 2012). Logo, há a necessidade de elaborar uma arquitetura viável para a execução dos cenários de testes, haja vista que esses são muito numerosos, na ordem de centenas;
- ✓ Análise e interpretação dos resultados do modelo adotado, fornecendo elementos necessários às conclusões consignadas na presente pesquisa, à luz do roteiro sugerido na pesquisas conduzida por Sun (2019).

Segundo Creswell (2007), existem três técnicas de pesquisa: Quantitativa, Qualitativa, ou Métodos Mistos. A técnica de pesquisa mais adequada para estudar as hipóteses propostas é a Quantitativa, na medida em que essa técnica emprega a coleta de dados e instrumentos estatísticos. Necessariamente, a utilização de dados comparativos, oriundos dos diferentes algoritmos de aprendizado de máquina supervisionados *Random Forest* desenvolvidos no ambiente computacional em linguagem *Python* com o propósito de testar as hipóteses, será um ponto chave para o desdobramento da pesquisa.

O maior risco inerente ao processo do sistema “*Malha Fina De Convênios*” é a classificação inadvertida de convênios cujas contas foram rejeitadas com uma *score* próximo de zero (falso positivo). Assim, o desempenho dos algoritmos de IA concebidos para testar as hipóteses será mensurado pela taxa de ocorrência de falsos positivos ao longo do domínio do *score* entre 0 e 1. Por seu turno, a precisão do algoritmo é aferida verificando-se a distribuição de convênios reprovados ao longo do intervalo [0, 1].

Conquanto os algoritmos de aprendizagem supervisionados *Random Forest* sejam de alto desempenho (Bao et al. 2020; Hooda, Bawa e Rana 2020; Breiman 2001; Grover, Bauhoff, Friedman 2019), isso tem pouco valor nos constructos propostos. O principal objetivo nos testes das hipóteses propostas nesta pesquisa é detectar com precisão o maior número possível de convênios reprovados, sem classificá-los erroneamente como aprovados, ou seja, a taxa de falso positivos em relação ao total. Nesse aspecto, é importante destacar uma peculiaridade em relação às métricas especificidade e sensibilidade, na medida em que essas métricas não são relativas ao total da população, mas às categorias convênios aprovados (sensibilidade) e reprovados (especificidade).

Ademais, as métricas para mensuração de desempenho geralmente utilizadas (acurácia, sensibilidade e especificidade) não são apropriadas para o enfrentamento deste problema de pesquisa haja vista o desequilíbrio no balanceamento da sequência de treinamento (Alpaydin 2020; Bao et al. 2020). Os convênios reprovados representam menos de 3% do total. Naturalmente, a estratégia de aferir a especificidade por meio da taxa de acerto de classificação de convênios reprovados levaria a uma precisão muito alta, o que seria ingênuo para avaliar o desempenho.

O desequilíbrio das sequências de treinamento, conforme retratado na literatura (Bao et al. 2020; Breiman 2001; Fawcett 2006; Domingos 2012; Stanisic, Radojevic, Stanic 2019), pode ser observado na Tabela 5, abaixo. Enquanto a extinta Secretaria de Desenvolvimento Agropecuário e Cooperativismo não rejeitou nenhuma prestação de conta dos 1.815 convênios em que essa Secretaria figurou como repassadora de recursos, o Ministério do Turismo reprovou 779 prestações de contas entre os 6.796 convênios celebrados por este Ministério na condição de repassador de recursos.

Tabela 5 – Distribuição da situação das prestações de contas dos convênios dos 10 órgãos mais representativos da população.

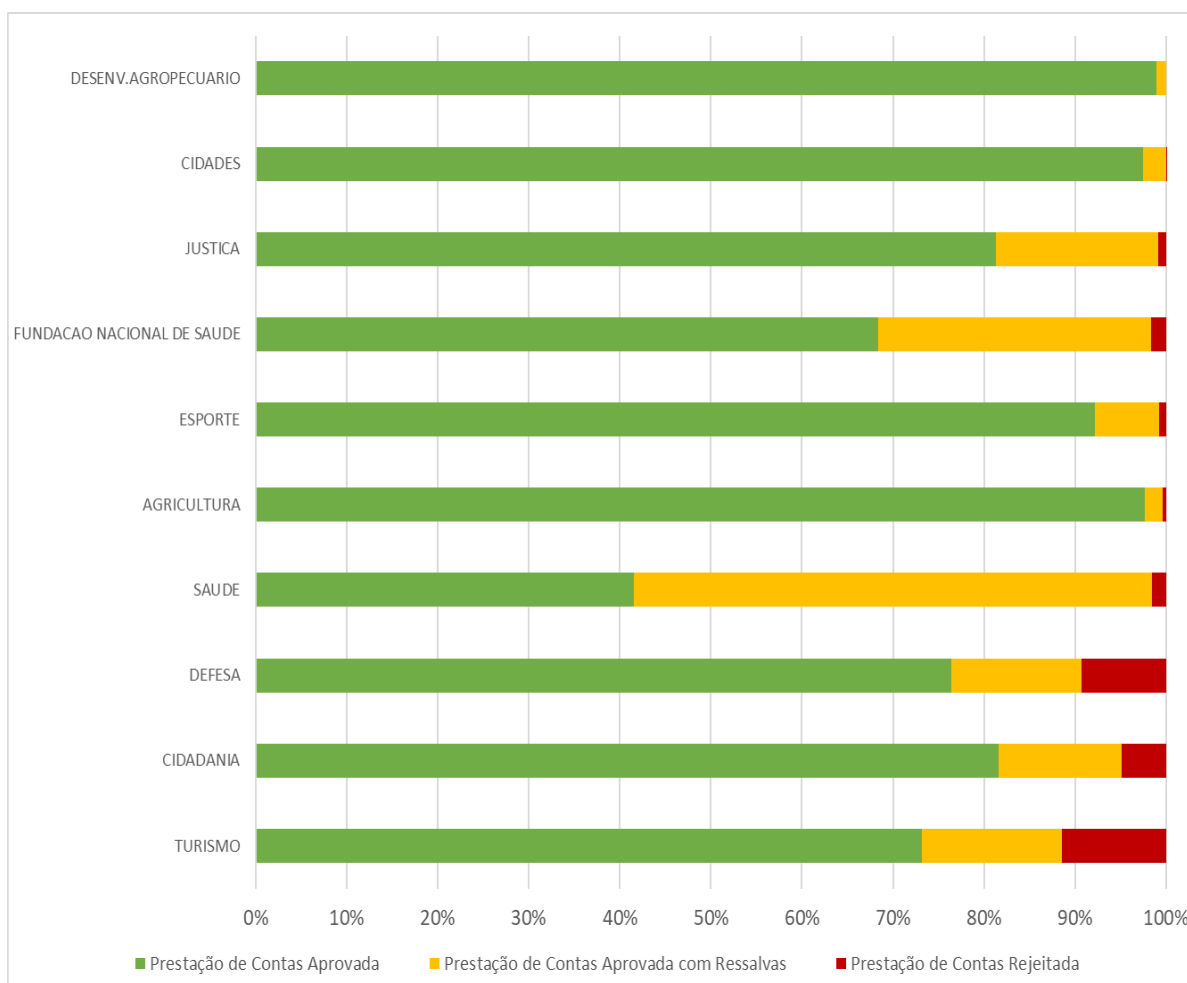
Descrição do Órgão	Prestações de Contas			Total
	Rejeitada	Aprovada	Aprovada com Ressalvas	
Secretaria de Desenvolvimento Agropecuário e Cooperativismo	0	1795	20	1815
Ministério das Cidades	2	9012	233	9247
Ministério da Justiça e Segurança Pública	7	688	151	846
Fundação Nacional de Saúde	16	677	297	990
Ministério do Esporte	29	3553	273	3855
Ministério da Agricultura, Pecuária e Abastecimento	46	12453	249	12748
Ministério da Saúde	84	2295	3148	5527
Ministério da Defesa	134	1104	206	1444
Ministério da Cidadania	137	2274	377	2788
Ministério do Turismo	779	4972	1045	6796
Total 10 órgãos	1234	38823	5999	46056
Total Todos os Órgãos	1533	43575	6737	51845

Fonte: Plataforma +Brasil. Data base: novembro de 2019.

Igualmente, a Figura 12 ilustra a diferença de rigor na análise das prestações de contas entre os 10 órgãos eleitos (Tabela 6) para o teste das hipóteses do constructo “*Identidade do repassador de recurso*”. Os Ministérios do Turismo, Cidadania e Defesa, nitidamente, rejeitam

mais prestações de contas do que a extinta Secretaria de Desenvolvimento Agropecuário e Cooperativismo e os Ministérios das Cidades e Justiça e Segurança Pública. Presume-se que os primeiros órgãos são mais austeros e comedidos nas análises de prestações de contas, enquanto os segundos são mais lenientes. Entretanto, frise-se que tal afirmação é meramente sugestiva, pois seria necessário avaliar o mérito de cada prestação de contas apresentada àqueles órgãos com baixa taxa de reprovação para embasar qualquer alegação sobre o zelo e diligência acerca da sua análise, afinal de contas as transferências de recursos aos entes subnacionais são cercadas de idiossincrasias (Amorim Neto e Simonassi 2013; Abrucio e Franzese 2007; Fajardo 2016). O fato de que os convênios celebrados nesses órgãos sejam conduzidos com a boa e regular aplicação dos recursos públicos não pode ser afastada, razão pela qual as contas de seus convênios nunca são rejeitadas.

Figura 12 - Distribuição percentual do resultado da análise das prestações de contas dos convênios dos 10 órgãos mais representativos da população



Logo, esta metodologia de pesquisa se valerá da métrica AUC, obtida pelo cômputo da área da curva ROC, além da proposição de uma nova métrica escalar derivada da acurácia (Alpaydin 2020), a qual denominamos inexatidão (ε). O valor da inexatidão (ε) de um algoritmo treinado por uma amostra de convênios em um determinado intervalo é:

$$\varepsilon_{ij} = \frac{FP_{ij}}{VP_{ij} + VN_{ij} + FN_{ij} + FP_{ij}}$$

No qual:

FP é Falso Positivo;

VP é Verdadeiro Positivo;

VN é Verdadeiro Negativo;

FN é Falso Negativo;

i: intervalos de score [0,0 a 0,1); [0,1 a 0,2); [0,2 a 0,3); [0,3 a 0,4); [0,4 a 0,5); [0,5 a 0,6); [0,6 a 0,7); [0,7 a 0,8); [0,8 a 0,9); e [0,9 a 1,0];

j: algoritmo treinado por uma determinada amostra de convênios, conforme a volumetria ou identidade do concedente;

Logo, a inexatidão (ε) de um algoritmo para uma faixa acumulada no intervalo de score é:

$$\varepsilon_j = \sum_0^i \frac{FP_{ij}}{VP_{ij} + VN_{ij} + FN_{ij} + FP_{ij}}$$

Adicionalmente, a métrica AUC para determinado algoritmo treinado por uma amostra de convênios é calculada pela integral da curva ROC, calculando a função da taxa de verdadeiros positivos em relação aos falsos positivos:

$$AUC_j = \int_0^1 \frac{VP_j}{VP_j + U - (VP_j + VN_j + FP_j)} dFP_j$$

No qual:

FP é Falso Positivo;

VP é Verdadeiro Positivo;

VN é Verdadeiro Negativo;

U é o Universo de Convênios, $U - (VP + VN + FP)$ é igual a FN;

j: algoritmo treinado por uma determinada amostra de convênios, conforme a volumetria ou identidade do concedente;

Encorajados por Cecchini et al. (2010), Fawcett (2006) e Grover, Bauhoff e Friedman (2019), a proposição da métrica inexatidão (ϵ) visa mensurar o real desempenho do algoritmo sobre as condições de testes impostas pelos constructos *Identidade do repassador de recurso* (H1 e H2) e *Tamanho de treinamento de dados* (H3 e H4) com foco na taxa de falsos positivos em relação ao total de convênios. Isso se deve porque a métrica especificidade reflete a proporção da categoria convênios aprovados, enquanto a métrica sensibilidade retrata a proporção da categoria convênios reprovados, ou seja, são métricas que abordam o desempenho isoladamente em subconjuntos do universo da população. Por seu turno, a métrica acurácia mede taxa de verdadeiros positivos e negativos conjuntamente sob as duas categorias de convênios aprovados e reprovados, tornando-a inapropriada, já que se pretende analisar o desempenho sob a ótica de falsos positivos em relação ao total.

Diante do exposto, o desempenho dos algoritmos gerados mediante a parametrização da identidade dos órgãos concedentes e da volumetria da sequência de treinamento será avaliada pelas métricas AUC e inexatidão (ϵ). A métrica AUC refletirá a qualidade intrínseca do algoritmo gerado em relação a sua sequência de treinamento (Larcker e Zakolyukina 2012) e a métrica inexatidão (ϵ) repercutirá a precisão do algoritmo quando esse classifica convênios diferentes daqueles usados em seu treinamento (Cecchini et al. 2010). Em todos os cenários desta pesquisa os dados de teste foram aqueles convênios com ciclo de vida encerrado na Plataforma +Brasil.

Por outro lado, cabe destacar que a tomada de decisão nesta pesquisa pelas métricas AUC e inexatidão (ϵ) para avaliar o desempenho dos algoritmos de *Machine Learning* se pautou em uma problemática recorrente do campo de pesquisa de Controle e Auditoria. Dentro da ciência social Administração (Fayol 1949; Taylor 1911), as pesquisas com enfoque na precisão de algoritmos de IA empreendem esforços para esclarecer a eficácia da substituição da intervenção humana, notadamente um auditor, por um mecanismo computacional para emitir a opinião sobre uma operação de negócio das organizações. Essas operações incluem os relatórios de demonstrações contábeis, regularidade de processos ou comprovação de despesas. Nesse

aspecto, Bao et al. (2020), Cecchini et al. (2010), Grover, Bauhoff e Friedman (2019), Fawcett (2006), Hooda, Bawa e Rana (2020), Larcker, Zakolyukina (2012), Stanisic, Radojevic e Stanic (2019) e Zhang(2019a) utilizam em suas pesquisas a métrica AUC e uma derivação da métrica inexatidão (ϵ) em detrimento de outras métricas utilizadas em pesquisas das ciências exatas como aquelas retratadas por Domingos (2012).

4.1. Relevância para o algoritmo de IA em distinguir a identidade dos repassadores de recursos

Evidentemente, cada órgão concedente possui seu próprio processo de análise de prestação de contas, com suas características e propriedades inerentes. Logo, ao universalizar o processo de aprendizagem de máquina indistintamente entre todos os repassadores de recursos, órgãos mais rigorosos na análise de prestação de contas podem sofrer uma atenuação indesejada na classificação de risco de seus convênios ao agregar órgãos menos rigorosos na análise.

O constructo “*Identidade do repassador de recurso*” da Tabela 4 foi concebido com o intuito de dirimir a dúvida erigida sobre a precisão (Bao et al. 2020; Zhang 2019) em circunstâncias diferentes quanto ao órgão concedente, uma vez que as métricas de desempenho do algoritmo foram estabelecidas uniformemente para todos os órgãos repassadores de recursos.

Nesses termos, o algoritmo de IA foi treinado separadamente para aqueles órgãos concedentes com a maior representatividade do universo dos convênios da Plataforma +Brasil cujo ciclo de vida encontra-se encerrado. Isso implica que a prestação de contas desses convênios foi analisada pelo órgão concedente, recebendo uma classificação como: a) prestação de contas aprovada; b) prestação de contas aprovada com ressalvas; ou c) prestação de contas rejeitada.

A população em estudo consiste em um quantitativo de 51.845 convênios que totalizam mais de R\$ 24 bilhões (R\$ 24.481.496.942,56). Na medida em que os convênios celebrados pelos órgãos concedentes da amostra da Tabela 6 espelham aproximadamente 90% do quantitativo e 80% de materialidade, a representatividade estatística dessa amostra em relação à população está satisfeita.

Tabela 6 - Amostra utilizada para os testes do constructo “*Identidade do repassador de recurso*”.

Órgão Concedente	Quantidade	Valor
-------------------------	-------------------	--------------

	#	% do Universo	R\$ Milhões	% do Universo
Ministério da Agricultura, Pecuária e Abastecimento	12828	24,74%	R\$ 4.103,58	16,76%
Ministério das Cidades	9306	17,95%	R\$ 3.632,21	14,84%
Ministério do Turismo	6838	13,19%	R\$ 2.997,27	12,24%
Ministério da Saúde	5642	10,88%	R\$ 2.616,30	10,69%
Ministério do Esporte	3883	7,49%	R\$ 1.307,82	5,34%
Ministério da Cidadania	2820	5,44%	R\$ 2.252,18	9,20%
Secretaria de Desenvolvimento Agropecuário e Cooperativismo	1816	3,50%	R\$ 466,27	1,90%
Ministério da Defesa	1447	2,79%	R\$ 849,52	3,47%
Fundação Nacional de Saúde	991	1,91%	R\$ 516,64	2,11%
Ministério da Justiça e Segurança Pública	861	1,66%	R\$ 967,54	3,95%
Total Amostra	46432	89,56%	R\$ 19.709,34	80,51%
Universo	51845	100%	R\$ 24.481,50	100%

Dessa forma, o algoritmo foi treinado separadamente com 10 séries de dados diferentes, uma para cada órgão concedente. Conseqüentemente, o mesmo procedimento utilizado para gerar o algoritmo do “*Malha Fina de Convênios*” será retomado, porém, com uma diferença substancial. Esse novo procedimento tem por objetivo treinar o algoritmo de aprendizado de máquina, exclusivamente, conforme a identidade do repassador dos recursos, ou seja, serão gerados 10 algoritmos diferentes, cada um refletindo a maneira que os órgãos repassadores de recursos analisam as prestações de contas apresentadas. Assim, haverá um balizador que permitirá o cotejamento das métricas AUC e inexatidão (ϵ) (Bao et al. 2020; Zhang 2019a;) entre a(i) a aplicação do algoritmo gerado por todo o universo de convênios em cada um dos órgãos repassadores de recursos e (ii) a aplicação do algoritmo gerado por convênios de um determinado órgão repassador em seus próprios convênios.

Essa estratégia permitirá avaliar o comportamento do aprendizado de máquina separadamente por órgão concedente. Pretende-se detectar se sequências de treinamento diferenciadas pela identidade do órgão concedente, produziram algoritmos com desempenhos

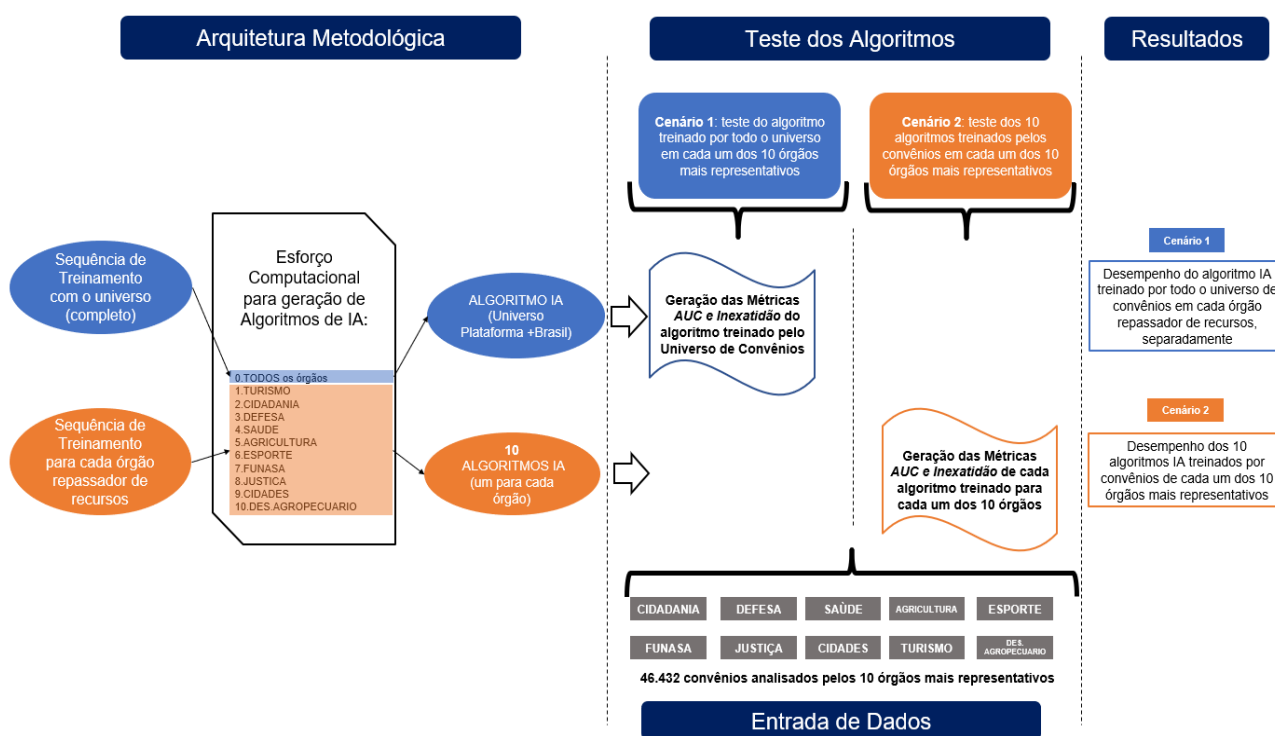
distintos a ponto de inviabilizar a extrapolação de um único algoritmo gerado a partir do aprendizado do universo.

A

Figura 13, abaixo, ilustra a estratégia para testar a hipótese do constructo “*Identidade do repassador de recurso*”. Dez algoritmos diferentes serão gerados conforme o órgão repassador de recursos para ser comparado com o algoritmo gerado por uma sequência de treinamento que contempla o universo de convênios registrados na Plataforma +Brasil.

Os resultados esperados eram a obtenção de dados que possibilitem a aferição da precisão dos algoritmos de IA por meio de duas métricas: inexatidão (ϵ) e AUC, conforme disciplinado em Bao et.al (2020), Breiman (2001), Fawcett (2006) e Zhang (2019a). Essas métricas foram utilizadas para aferir a precisão em duas configurações envolvendo a identidade dos órgãos concedentes. A primeira configuração, cenário 1, é aquela decorrente da aplicação do algoritmo gerado por todo o universo de convênios em cada um dos órgãos repassadores de recursos. A segunda, cenário 2, é aquela em que o algoritmo gerado por convênios de um determinado órgão repassador foi aplicado para classificar seus próprios convênios (Tabela 6).

Figura 13 - Metodologia para verificar o impacto da identidade do repassador de recursos na precisão(métricas AUC e inexatidão) do algoritmo de inteligência artificial.



Elaborado pelo autor.

4.2. Relevância do tamanho da série de treinamento para o algoritmo IA

O constructo “*Tamanho de treinamento de dados*” (Tabela 4) foi estabelecido com o intuito de averiguar o impacto que a volumetria da série de treinamento possui na precisão do algoritmo de aprendizado de máquina. O cerne da intervenção proposta está na verificação do comportamento da precisão de diferentes algoritmos, os quais foram gerados por sequências de treinamento incrementais, segmentadas por ano de celebração dos convênios. O comportamento dos resultados dos algoritmos será analisado à medida que mais convênios servem de insumo para o aprendizado de máquina (Breiman 2001; Zhang 2019).

No aprendizado supervisionado, os algoritmos precisam ser treinados primeiramente no conjunto de dados de treinamento e depois testados no conjunto de dados de teste. A regra geral na divisão entre conjunto de treinamento e teste é escolher cerca de 75% da amostra como dados de treinamento e o restante como dados de teste (Alpaydin, 2020; Hooda, Bawa e Rana 2020; Zhang 2019a;). Todavia, isso não é uma regra consagrada, uma vez que a garantia de desempenho do aprendizado de máquina depende da disponibilidade de rotulação dos dados (Sun 2019). Outra questão importante, conforme sublinha Cecchini et al. (2010), Grover, Bauhoff, Friedman (2019) e Zhang (2019a) é como garantir o desempenho do aprendizado de máquina, dada a disponibilidade e o balanceamento de dados que estão rotulados, no caso em epígrafe contas aprovadas ou rejeitadas. Nesse contexto, as potencialidades do uso do Malha Fina de Convênios, seus impactos já demonstrados e suas limitações ainda não são claros e definitivos. Logo, a motivação para a proposição do constructo “*Tamanho de treinamento de dados*” visa a compreensão desse cenário.

Em janeiro de 2020, a Plataforma +Brasil contava com 175.456 convênios cadastrados. Desses, 53.002 completaram seu ciclo de vida, ou seja, o órgão repassador de recursos avaliou a prestação de contas apresentada pelo conveniente definitivamente, aprovando-a, aprovando-a com ressalvas, ou rejeitando-a, conforme exibido na Tabela 7. A essência de avaliação desse constructo está na verificação do comportamento da precisão do algoritmo do sistema “*Malha Fina de Convênios*” ao longo do tempo.

Dessa forma, o algoritmo foi treinado com diferentes segmentos de dados conforme o ano de celebração do convênio, admitindo-se, dessa forma, 9 séries de dados diferentes entre os anos 2008 a 2016. Essas séries serão analisadas especificamente quanto ao comportamento da precisão do algoritmo.

Adicionalmente, é importante delimitar que a sequência de treinamento do algoritmo deve corresponder apenas aos convênios cujo seu ciclo de vida foi encerrado, ou seja, perpassou pelas fases da celebração, execução e prestação de contas. Consequentemente, isso exige uma fonte fidedigna e estável da base de dados da Plataforma +Brasil, que reflita uma perenidade das situações dos convênios. Na medida em que o processo das transferências voluntárias da união é inerentemente volátil (Brollo e Nannicini 2012; Meireles, 2019), alterações nas situações dos convênios podem ocorrer, principalmente, naqueles celebrados recentemente. Por conseguinte, adotou-se o ano de 2017 como ponto de corte da população de convênios a serem usados como sequência de treinamento.

Tabela 7 - Amostra utilizada para realizar os testes do construto “*Tamanho de treinamento de dados*”. Data base janeiro de 2020.

Ano	Qtd Acumulado Convênios	Prestação de Contas					
		Aprovada	Aprovada com Ressalvas	Rejeitada	Total	Acumulado	% Acumulado
2008	2746	797	165	182	1144	1144	2,16%
2009	26643	10515	1711	631	12857	14001	26,42%
2010	50186	9918	1495	462	11875	25876	48,82%
2011	63833	5855	1038	84	6977	32853	61,98%
2012	76465	4727	635	40	5402	38255	72,18%
2013	92712	6044	562	67	6673	44928	84,77%
2014	106225	3491	719	30	4240	49168	92,77%
2015	116386	1541	313	19	1873	51041	96,30%
2016	131413	1535	133	9	1677	52718	99,46%
2017	151931	200	22	1	223	52941	99,88%
2018	170683	55	6		61	53002	100,00%
2019	175418	0	0	0	0	53002	100,00%
Total	175418	44678	6799	1525	53002	53002	100,00%

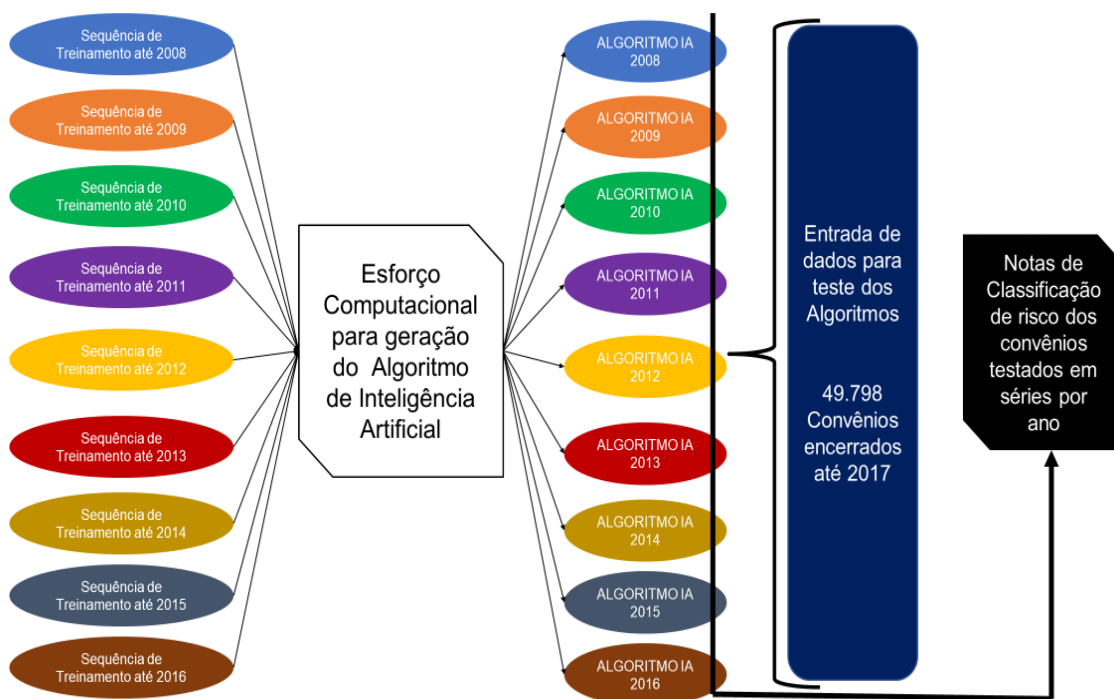
Fonte: Plataforma +Brasil.

Entretanto, mediante a observação da Tabela 7, percebe-se que após 2017 apenas 284 convênios tiveram seus ciclos de vida estão encerrados. Por sua vez, esse baixo quantitativo não possui significância para a sequência de treinamento. Esse fato acarretou a delimitação do escopo em 9 séries entres os anos de 2008 e 2016, excluindo a série de 2017.

A Figura 14 ilustra o processo realizado para determinar diferentes algoritmos de inteligência artificial com o incremento gradual das sequências de treinamento. As 9 séries utilizadas como sequência de treinamento produziram 9 algoritmos distintos entre si, capazes de classificar qualquer convênio da Plataforma +Brasil quanto à probabilidade de as contas

serem aprovadas ou rejeitadas. Por seu turno, os 9 algoritmos gerados foram testados mediante a análise das métricas AUC e inexatidão (ϵ) (Bao et al. 2020; Fawcett 2006; Zhang 2019a).

Figura 14 – Metodologia para verificar o impacto da volumetria na sequência de treinamento na precisão (métricas AUC e inexatidão) do algoritmo de IA.



Elaborado pelo autor.

Por fim, o teste dos algoritmos só poderia ser executado com os convênios cujo ciclo de vida se encerrou, analogamente à população usada para a sequência de treinamento. Ora, a forma mais assertiva de testar a precisão de algoritmos de aprendizado de máquina é a comparação com o resultado real do convênio. Da mesma forma, o algoritmo aprende com os convênios encerrados, ou seja, com as variáveis que os levaram a ter suas contas rejeitadas ou encerradas. Assim, o teste dos 9 algoritmos consistiu em classificar a completude dos convênios da Tabela 7.

5. ANÁLISE DOS RESULTADOS

A seguir são apresentados os resultados obtidos e a sua análise. Os resultados consistem nas métricas *AUC* e *inexatidão* (ϵ) geradas nos diferentes cenários de aprendizado de máquina, os quais foram concebidos por meio da parametrização da identidade do repassador de recursos e da volumetria da sequência de treinamento.

5.1.1. Métricas *AUC* e *inexatidão* (ϵ) conforme a identidade do órgão concedente

A avaliação do desempenho da precisão dos algoritmos treinados conforme a identidade do órgão concedente sob a perspectiva da métrica *AUC* permite concluir que alguns possuem desempenho melhor do que o algoritmo treinado com a população completa dos convênios. Entretanto, a avaliação sob a perspectiva da métrica *inexatidão* (ϵ) aponta que nenhum algoritmo possui desempenho melhor do que aquele treinado pelo universo. Na medida em que a superação de desempenho sob o ponto de vista da métrica *AUC* ocorre em apenas 4 algoritmos, de maneira residual, e sob o ônus de resultar em 6 algoritmos com desempenho inferior, sendo 2 desses muito insatisfatórios, a *H1* do constructo *Identidade do repassador de recurso* não é confirmada sob certas circunstâncias, as quais não ensejam a adoção da estratégia de treinar algoritmos separadamente conforme o órgão repassador de recursos. Por seu turno, a *H2* é confirmada porque o desempenho é diferente conforme a identidade do órgão concedente.

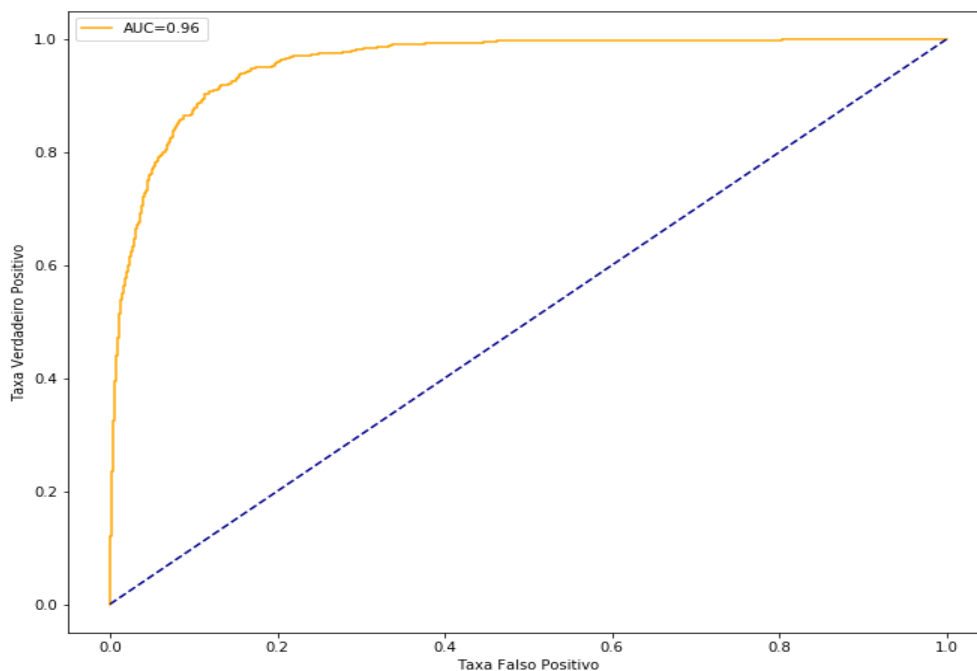
5.1.1.1. Cenário 1: algoritmo de IA treinado com o universo e aplicado em cada órgão concedente separadamente

O desempenho do algoritmo treinado a partir do universo é satisfatório sob o ponto de vista da métrica *AUC*, uma vez que assume o valor 0,96, conforme demonstrado na Figura 15. Essa inferência é corroborada por Fawcett (2006) e Zhang (2019a), porque esses afirmam que quanto mais a *AUC* se aproxima de 1, melhor é o desempenho de um algoritmo classificador. Entretanto, a análise da métrica *inexatidão* (ϵ) indica que o algoritmo possui desempenho diferente ao longo do intervalo [0 a 1], na medida em que ocorre aumento de probabilidade de falsos positivos nos órgãos cujo histórico de reprovação de contas de convênios é maior que a média⁷, conforme demonstrado na Tabela 8. Por outro lado, ocorre o inverso nos órgãos com histórico de reprovação de contas abaixo da média, pois há diminuição de falsos positivos. Não

⁷ A média de rejeição de prestação de contas da amostra dos 10 órgãos da **Tabela 6** (Ministério do Turismo, Ministério da Cidadania, Ministério da Defesa, Ministério da Saúde, Ministério da Agricultura, Ministério do Esporte, Funasa, Ministério da Justiça, Ministério das Cidades, Desenvolvimento Agropecuário) é 137,111.

obstante, a métrica inexatidão (ϵ) apresenta discrepância apenas a partir do intervalo (0,8 a 1].

Figura 15 - Curva ROC e métrica AUC do algoritmo de IA treinado com o universo.



A constatação de que a métrica inexatidão (ϵ) apresenta valores deficientes apenas a partir do intervalo (0,8 a 1] é inferida por meio da análise das curvas da Figura 16, gerada a partir da Tabela 8. Observa-se que a inexatidão (ϵ) apresenta valores muito bons entre o intervalo [0 a 0,8). Porém, a discrepância da inexatidão (ϵ) observada a partir do score 0,8 decorre da superestimação e da subestimação de falsos positivos em diferentes identidades dos órgãos concedentes. As curvas “DEFESA”, “TURISMO” e “CIDADANIA”, associadas aos convênios celebrados pelos Ministérios da Defesa, Turismo e Cidadania, respectivamente, estão notadamente superiores à curva “Todos os Órgãos” a partir do intervalo (0,8 a 1]. Infere-se, portanto, que o comportamento discrepante das curvas da métrica imprecisão, quando o algoritmo é aplicado nesses 3 órgãos, decorre da maior rejeição de contas de convênios promovidas por esses órgãos em relação à média, fazendo com que a sequência de treinamento seja menos desbalanceada. Essa conclusão vai ao encontro do que preconiza Alpaydin (2020) e os achados das pesquisas de Bao et al. (2020) e Cecchini et al. (2010).

Resumidamente, a análise dos resultados empíricos da métrica inexatidão (ϵ) permite concluir que o algoritmo treinado a partir da população completa da Plataforma +Brasil, quando aplicado separadamente em cada órgão concedente, apresenta desempenho inferior para os Ministérios da Defesa, Turismo e Cidadania a partir do score 0,8, com acentuação da degradação do desempenho a partir do score 0,9. Adicionalmente, os resultados indicam que o desempenho do algoritmo nos convênios do Ministério da Cidadania apresenta um início de degradação moderado a partir do score 0,8. Essa observação implica que a probabilidade de ocorrência de falsos positivos nesse Ministério possui um espectro mais abrangente do que os Ministérios da Defesa e Turismo.

Ademais, existe uma discrepância a partir do score 0,9, implicando em duas perspectivas. A primeira delas é que a inexatidão (ϵ) de 7 órgãos está subestimada em torno de -2%, significando um desvio moderado que não gera consequências adversas, na medida em que acarretará uma imposição de conservadorismo a esses órgãos na análise das contas de seus convênios. Ora, a subestimação de risco consiste na indicação pelo algoritmo de IA de mais falsos positivos, ou seja, mais convênios seriam reprovados do que realmente são. Contudo, a superestimação de erro gera consequências adversas, haja vista que ocorre a classificação de convênios reprováveis como possivelmente aprovados pelo algoritmo. Essa subestimação ocorre para os Ministérios da Defesa, Turismo e Cidadania, atingindo taxas de divergência aproximadas de 6%, 8% e 2%, respectivamente.

Órgão Concedente	Prestação de Contas Rejeitadas	Intervalo Acumulado de <i>Score</i>						
		[0,0 a 0,4)	[0,0 a 0,5)	[0,0 a 0,6)	[0,0 a 0,7)	[0,0 a 0,8)	[0,0 a 0,9)	[0,0 a 1,0]
Todos os Órgãos	1490	0,00000%	0,00195%	0,02140%	0,03890%	0,10310%	0,27428%	2,89838%
Ministério do Turismo	779	0,00000%	0,01471%	0,04414%	0,07357%	0,17657%	0,44144%	11,46263%
Ministério da Cidadania	137	0,00000%	0,00000%	0,03587%	0,03587%	0,17934%	0,86083%	4,91392%
Ministério da Defesa	134	0,00000%	0,00000%	0,00000%	0,00000%	0,00000%	0,27701%	9,27978%
Ministério da Saúde	84	0,00000%	0,00000%	0,01809%	0,03619%	0,05428%	0,21712%	1,51981%
Ministério da Agricultura	46	0,00000%	0,00000%	0,00000%	0,01569%	0,04707%	0,07844%	0,36084%
Ministério do Esporte	29	0,00000%	0,00000%	0,00000%	0,02594%	0,05188%	0,15564%	0,75227%
Funasa	16	0,00000%	0,00000%	0,20202%	0,20202%	0,20202%	0,30303%	1,61616%
Ministério da Justiça	7	0,00000%	0,00000%	0,11820%	0,11820%	0,11820%	0,23641%	0,82742%
Ministério das Cidades	2	0,00000%	0,00000%	0,01081%	0,01081%	0,01081%	0,01081%	0,02163%
Desenvolvimento Agropecuário	0	0,00000%	0,00000%	0,00000%	0,00000%	0,00000%	0,00000%	0,00000%

 Tabela 8 – Métrica Inexatidão (ϵ) do Algoritmo de IA treinado a partir do universo (Probabilidade de Falso Positivo).

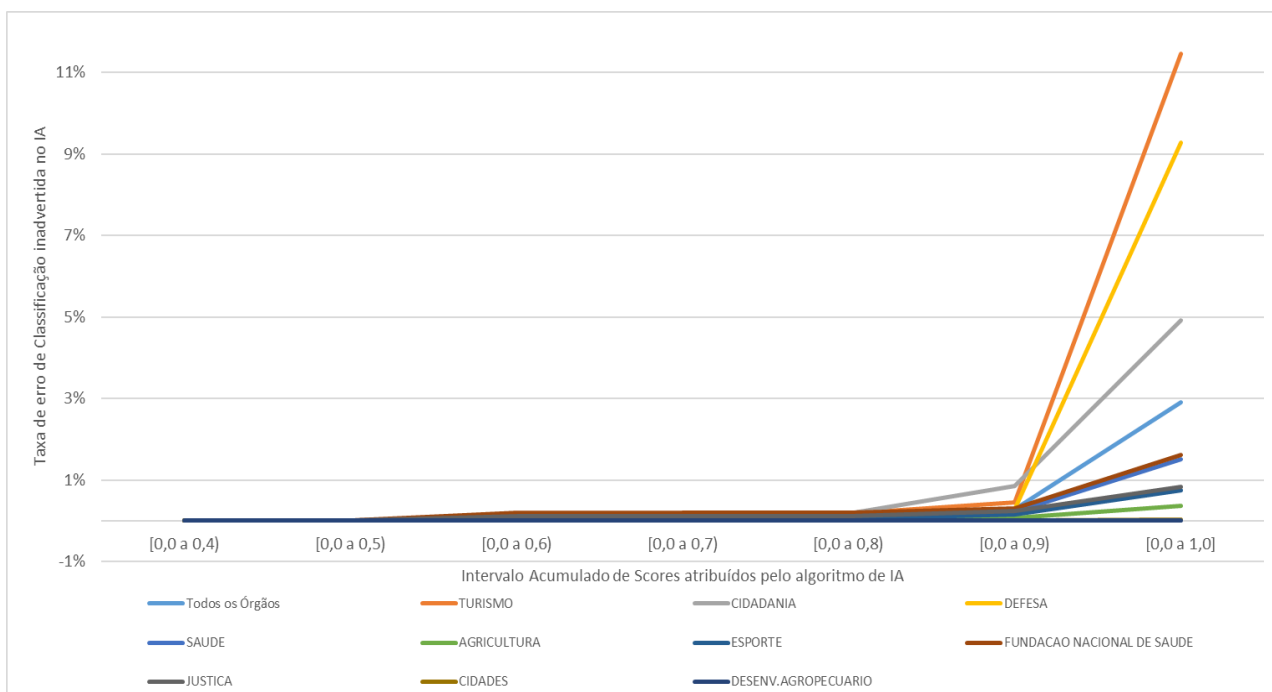


Figura 16 – Métrica Inexatidão (ϵ) do Algoritmo de IA treinado a partir do universo (Probabilidade de Falso Positivo).

5.1.1.2. Cenário 2: algoritmo de IA treinado com convênios conforme a identidade do órgão concedente

Os desempenhos dos algoritmos treinados a partir dos convênios de cada órgão concedente apresentam comportamentos distintos sob a perspectiva da métrica AUC, alguns apresentando desempenho muito satisfatório, enquanto outros bastante insatisfatórios. Considerando o algoritmo treinado a partir do universo como *benchmarking*, cujo AUC é 0,96, os algoritmos treinados por convênios dos Ministérios do Turismo, Agricultura, Esporte e Cidades, cujas métricas AUC assumiram 0,97, 0,97, 0,99 e 0,98, respectivamente (Figura 17), obtiveram desempenho muito satisfatório. A superioridade de desempenho do algoritmo treinado pelos convênios do Ministério do Turismo se deve pela melhor qualidade no balanceamento da sequência de treinamento entre todos os Ministérios, corroborando com os achados das pesquisas conduzidas por Bao et. al (2020) e Zhang (2019a). Por seu turno, o desempenho superior dos algoritmos treinados pelos convênios dos Ministérios da Agricultura e Cidades se justificam pela volumetria da sequência de treinamento, na medida em que são os

dois órgãos concedentes com maior quantidade de prestação de contas analisadas, 24,74% e 17,95% respectivamente, indo ao encontro do que afirmam Cecchini et al. (2010) e Larcker e Zakolyukina (2012). Por outro lado, o desempenho quase ideal do algoritmo do Ministério do Esporte, com AUC assumindo o valor 0,99, não possui embasamento na volumetria ou no equilíbrio da sequência de treinamento, já que esse Ministério é apenas o quinto em volumetria, com 7,49% do total, e a quantidade de prestação de contas rejeitada, 29, é bem abaixo da média de 137. Essa singularidade é explicada por Alpaydin (2020), Breiman (2001) e Domingos (2012) quando alguma característica da sequência de treinamento permite a identificação inequívoca das instâncias a serem classificadas, fazendo com que os falsos positivos e falsos negativos sejam reduzidos a quase zero.

Ademais, a Figura 17 ilustra o desempenho dos algoritmos treinados pelos convênios dos Ministérios da Cidadania, Defesa e Saúde, cujas métricas AUC assumiram respectivamente 0,91, 0,91 e 0,92. Conforme sublinha Fawcett (2006), seu desempenho é satisfatório em virtude de a AUC ser maior que 0,9, mas os dados empíricos obtidos nesta pesquisa demonstram que seu desempenho não superou aquele obtido no algoritmo treinado pelo universo ($AUC=0,96$). Em contrapartida, os algoritmos da Funasa e do Ministério da Justiça obtiveram desempenho muito insatisfatório, pois sua AUC foram 0,63 e 0,77, respectivamente. Esses resultados endossam as conclusões de Bao et al. (2020) e Zhang (2019a), na medida em que a volumetria dos convênios desses dois órgãos são as duas menores entre todos os órgãos (1,91%, Funasa, 1,66%, Ministério da Justiça) e o balanceamento é muito desequilibrado (16 reprovações para a Funasa e 7 para o Ministério da Justiça). Por sua vez, não houve possibilidade de produzir um algoritmo para a Secretaria de Desenvolvimento Agrário na medida em que esse órgão nunca reprovou uma prestação de contas. Conseqüentemente, não há métricas para a mensuração de desempenho para esse órgão.

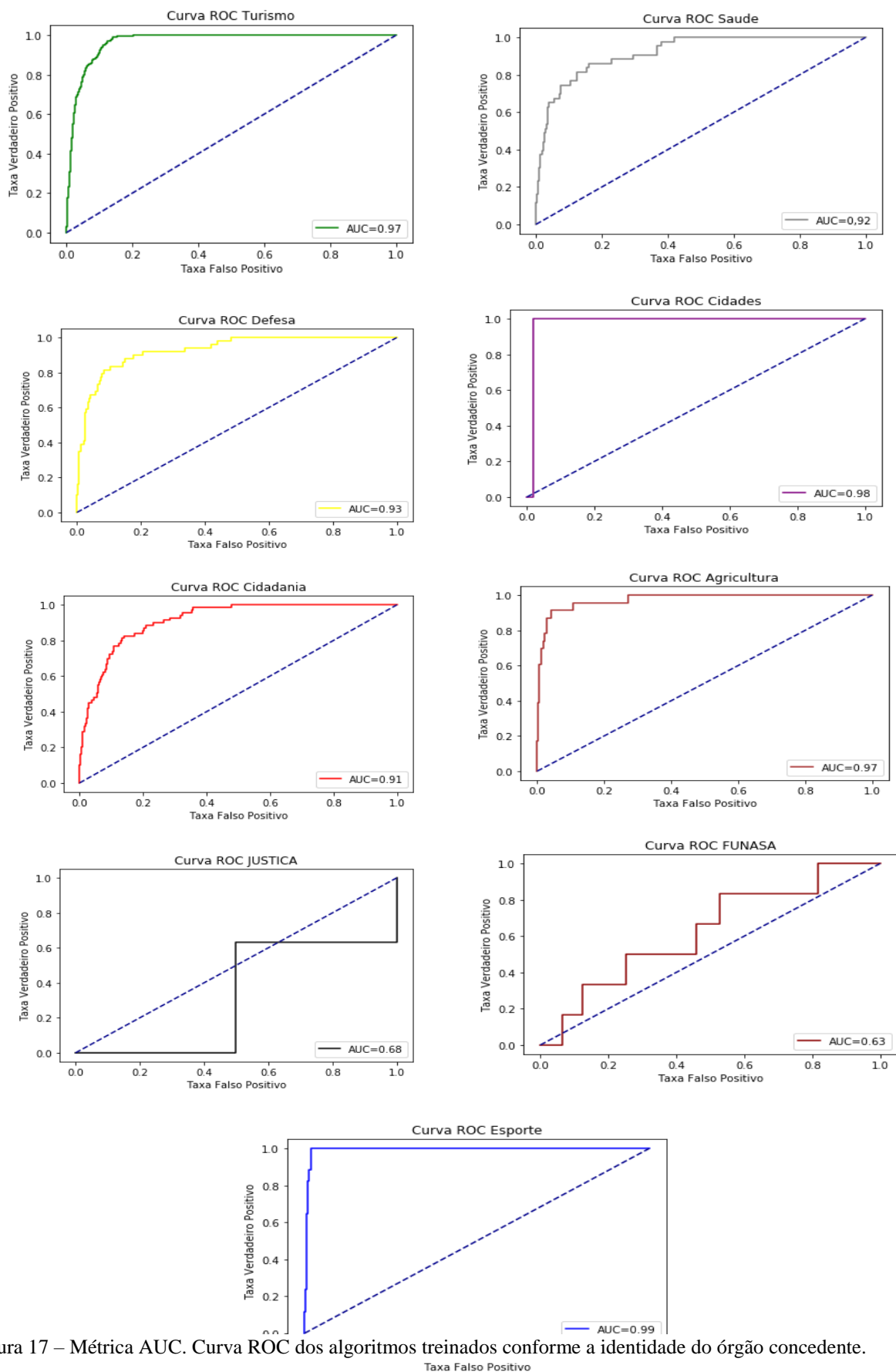
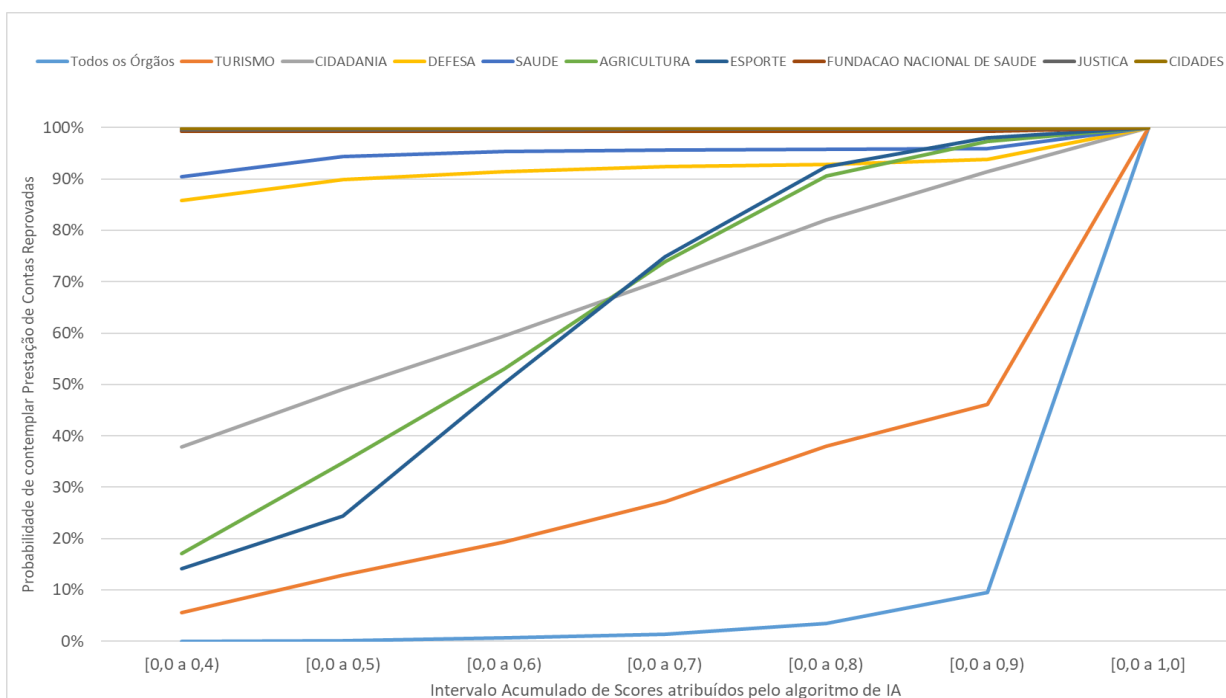


Figura 17 – Métrica AUC. Curva ROC dos algoritmos treinados conforme a identidade do órgão concedente.

Taxa Falso Positivo

Essa constatação é evidenciada por meio da Figura 18, a seguir. A análise a ser feita reside na comparação da métrica inexactidão (ϵ) entre o (a) algoritmo treinado a partir da população completa, curva “Todos os órgãos”, e (b) os algoritmos treinados a partir dos convênios de cada órgão concedente. A Figura 18 ilustra o comportamento da métrica inexactidão (ϵ) entre os 11 algoritmos treinados, oportunidade em que é possível observar a quantidade de convênios categorizados em cada intervalo acumulado de score entre [0 a 1]. Assim, o algoritmo que consegue concentrar uma quantidade maior de convênios em intervalos acumulados mais próximos de 1 possuem desempenho melhor. Infere-se, portanto, que o algoritmo representado pela curva “Todos os órgãos” é o que possui a melhor métrica inexactidão (ϵ).

Figura 18 – Curvas da Métrica Inexactidão (ϵ) dos Algoritmos de IA treinados a partir dos convênios de cada órgão concedente.



Adicionalmente, observa-se, sob o ponto de vista da métrica inexactidão (ϵ), que o algoritmo treinado pelos convênios do Ministério do Turismo, curva ”Turismo”, possui desempenho melhor que os demais algoritmos, entretanto sem superar aquele treinado com a população completa. Isso se deve ao fato de o Ministério do Turismo ter o maior histórico de reprovação de contas de convênios entre todos os órgãos, 779 convênios reprovados entre 2008 e 2017, fazendo com que a sequência de treinamento seja a menos desequilibrada (Bao et al. 2020; Zhang 2019a). Os algoritmos estampados nas curvas “Cidades” e “Desenvolvimento

Agrário” endossam esse argumento. As sequências de treinamento desses algoritmos possuem, respectivamente, 2 e 0 convênios com prestação de contas rejeitadas, proporcionando sequências de treinamento com intenso desequilíbrio e desempenho ruim.

Não obstante, destaca-se que uma sequência de treinamento equilibrada é condição necessária, mas não suficiente para prover bom desempenho. Observa-se que o desempenho da curva ”Turismo” sobressai em virtude do desequilíbrio moderado da sequência de treinamento aliado a sua alta volumetria, uma vez que a quantidade relativa de convênios com prestação de contas concluída no Ministério do Turismo é de 13,19% do total (Tabela 6), fazendo com que seja a terceira maior. Isso reforça as alegações de Alpaydin (2020), Breiman (2001) e Domingos (2012) que sequências de treinamento maiores proporcionam melhores desempenho.

A ambiguidade entre volumetria e balanceamento da sequência de treinamento é retratada nas curvas “Agricultura”, “Esporte” e “Cidadania”. Conquanto o Ministério da Agricultura tenha a maior quantidade de convênios celebrados entre todos os órgãos (24,74% do total), isso não foi suficiente para gerar um algoritmo com bom desempenho. O Ministério da Agricultura carece de equilíbrio em sua sequência de treinamento, uma vez que possui apenas 46 convênios com prestações de contas rejeitadas. A mesma situação ocorre com o Ministério do Esporte, que possui uma grande volumetria relativa (7,49 % do total) da sequência de treinamento, mas também um severo desbalanceamento pois reprovou apenas 29 convênios.

Por outro lado, observa-se algumas situações em que os algoritmos possuem baixo desempenho, embora tenham sido gerados com sequências de treinamento relativamente balanceadas. Nessas situações, o baixo desempenho da métrica inexatidão (ϵ) é provocado pela baixa volumetria da sequência de treinamento. O algoritmo representado pela curva “Defesa” foi gerado com uma sequência de 134 convênios reprovados, proporcionado um balanceamento relativamente alto, mas a volumetria representa apenas 2,79% do total.

Outrossim, há aqueles algoritmos com desempenho bastante degradado porque suas sequências de treinamento congregam nuances de baixa volumetria e alto desequilíbrio. Os algoritmos representados pelas curvas “Funasa”, “Saude”, “Justica” inserem-se nesse contexto. Por sua vez, não houve possibilidade de produzir um algoritmo para a Secretaria de Desenvolvimento Agrário na medida em que esse órgão nunca reprovou uma prestação de contas.

5.1.2. Métricas *AUC* e *inexatidão* (ϵ) conforme a volumetria da sequência de treinamento.

A avaliação do desempenho da precisão dos algoritmos treinados mediante o incremento gradual da sequência de treinamento, sob a perspectiva da métrica *AUC*, permite concluir que o desempenho melhora à medida que aumenta a volumetria de dados para treinamento. Todavia, a melhoria atinge um patamar em que qualquer incremento no tamanho da sequência de treinamento não implica em melhoria de desempenho. Da mesma forma, a avaliação, sob o ponto de vista da métrica *inexatidão* (ϵ), permite inferir as mesmas conclusões obtidas sob o prisma da métrica *AUC*: quanto mais volumetria, melhor o desempenho, e existe um limite em que o incremento na volumetria não produz melhoria no desempenho. Logo, a *H3* do constructo *Tamanho de treinamento de dados* é confirmada, assim o a *H4*, que versa acerca do ponto de saturação, não obstante a saturação na métrica *AUC* (ano de 2012) ter sido ligeiramente diferente da métrica *inexatidão* (ϵ) (ano de 2014).

Sob a perspectiva da métrica *AUC*, conforme ilustrado na Figura 19, os desempenhos dos algoritmos treinados a partir do incremento gradual do tamanho da sequência de treinamento apresentam comportamento de melhoria contínua até atingirem o mesmo patamar de desempenho do algoritmo treinado pelo universo, esse último convencionado como *benchmarking*, cujo *AUC* assumiu o valor de 0,96. Esse patamar limítrofe foi obtido com o algoritmo treinado com os convênios celebrados até o ano de 2012, na medida em que a *AUC* dessa instância assumiu o valor 0,96. A partir do algoritmo de 2012, todos os subsequentes obtiveram o mesmo desempenho, ou seja, ocorreu um ponto de saturação quando a sequência de treinamento atingiu aproximadamente 72% do tamanho da população (Tabela 7). Esse achado vai ao encontro do que alegam Alpaydin (2020) e Zhang(2019a) sobre o tamanho ideal de 75% do total da população a ser utilizada como amostra para treinamento do algoritmo.

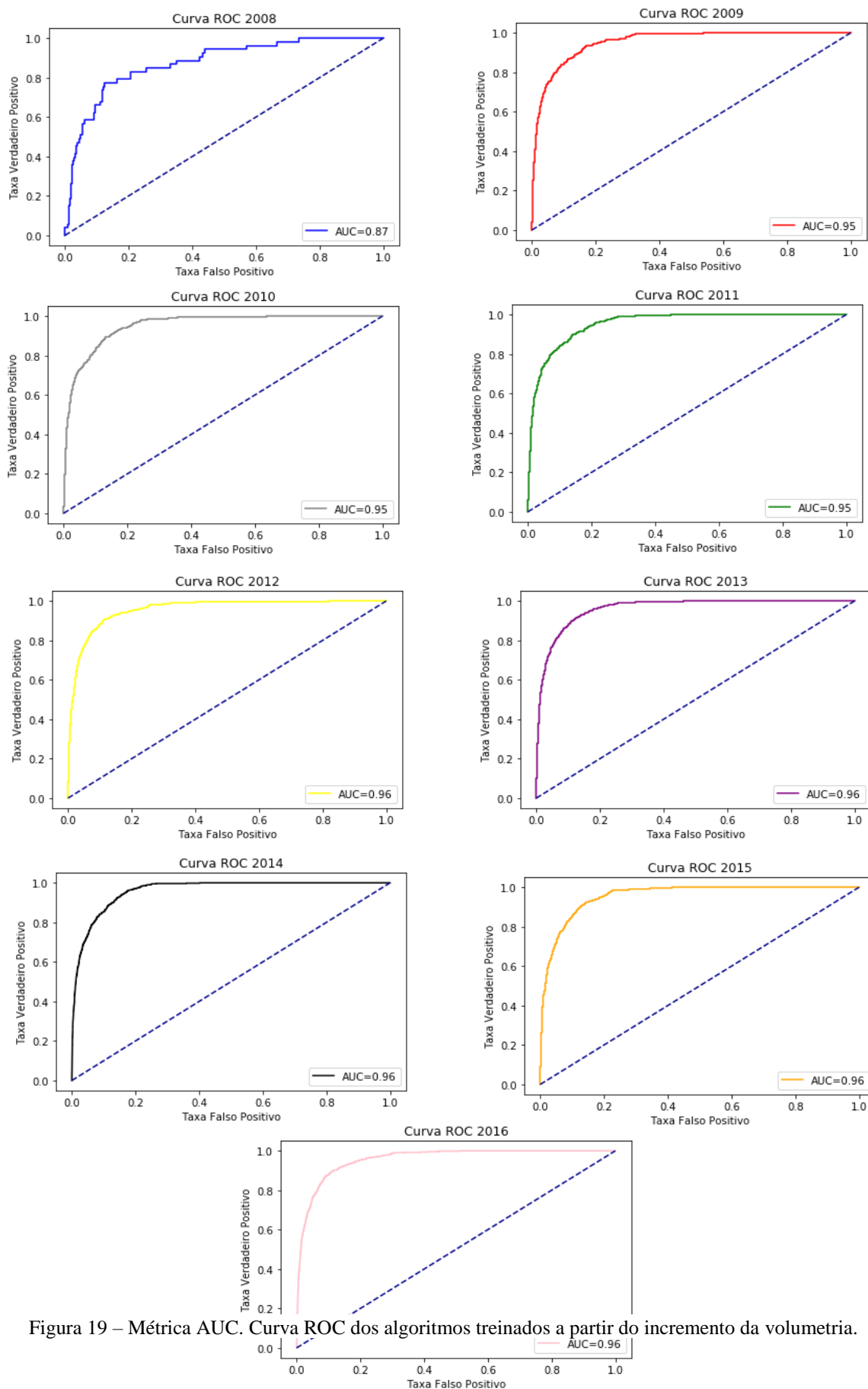


Figura 19 – Métrica AUC. Curva ROC dos algoritmos treinados a partir do incremento da volumetria.

Por outro lado, a Tabela 10 exibe o desempenho sob a perspectiva da métrica inexatidão (ϵ) mediante a variação da volumetria da sequência de treinamento, incrementada sequencialmente a cada ano a partir de 2008. A observação dessa tabela permite inferir que o algoritmo treinado com os convênios celebrados em 2008 possui um desempenho muito insatisfatório sob a perspectiva da inexatidão (ϵ).

Tabela 10 – Métrica Inexatidão (ϵ) dos Algoritmos de IA treinados a partir dos convênios separados por ano de celebração. (Probabilidade de Falso Positivo).

Ano Alcance	Intervalo Acumulado de <i>Score</i>						
	[0,0 a 0,4)	[0,0 a 0,5)	[0,0 a 0,6)	[0,0 a 0,7)	[0,0 a 0,8)	[0,0 a 0,9)	[0,0 a 1,0]
2008	85,55%	87,05%	88,26%	89,61%	90,68%	91,74%	100,00%
2009	0,93%	3,56%	12,53%	24,98%	36,01%	45,34%	100,00%
2010	0,07%	0,93%	2,21%	8,75%	17,94%	22,28%	100,00%
2011	0,00%	0,93%	3,56%	9,61%	11,74%	17,37%	100,00%
2012	0,00%	0,21%	1,14%	4,27%	9,82%	16,73%	100,00%
2013	0,00%	0,14%	1,28%	2,57%	5,49%	11,90%	100,00%
2014	0,00%	0,14%	0,36%	1,85%	4,13%	9,54%	100,00%
2015	0,00%	0,14%	0,57%	1,71%	4,20%	11,32%	100,00%
2016	0,00%	0,21%	0,57%	1,28%	4,13%	9,40%	100,00%

Seguidamente, o algoritmo treinado com os convênios celebrados até 2009 produziu um modelo melhor que o anterior, porém ainda insatisfatório porque a métrica inexatidão (ϵ) continua muito degradada. Ato contínuo, infere-se que o algoritmo treinado pelos celebrados convênios até 2010 é pior que o de 2011, que por sua vez é pior que o de 2012, que por sua vez é pior que o de 2013. Consequentemente, essa inferência endossa a hipótese suscitada de que o desempenho do algoritmo de aprendizado de máquina aumenta à medida que há mais dados no universo de aprendizagem, o que vai ao encontro da literatura recente (Bao et al. 2020; Cecchini et al. 2010; Larcker e Zakolyukina, 2012; Hooda, Bawa e Rana 2020). Todavia, merece destaque o fato de que o algoritmo de 2010 é melhor que o algoritmo de 2011 até o intervalo acumulado [0 a 0,6), apesar de este algoritmo ter sido gerado com uma volumetria de sequência de treinamento maior que aquele, o que de certa forma contraria o que assevera Alpaydin (2020). No entanto, Breiman (2001) assevera que essas situações que escapam à dicotomia mais volumetria e melhor desempenho podem ocorrer quando alguma característica da sequência de

treinamento faz com que o algoritmo identifique as classes explicitamente, nesse caso convênios reprovados e aprovados.

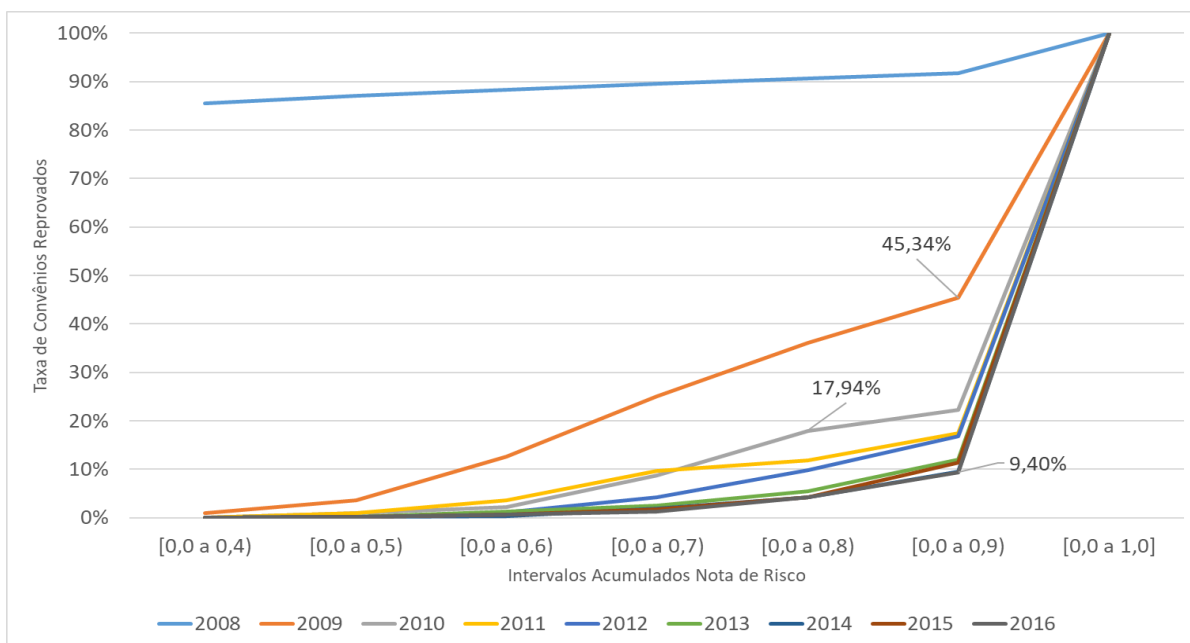


Figura 20 - Curva da Métrica Inexatidão (ϵ) dos Algoritmos de IA treinados a partir dos convênios separados por ano de celebração. (Probabilidade de Falso Positivo).

Outrossim, infere-se que o aprimoramento do desempenho dos algoritmos treinados com volumetria incremental a cada ano, possui uma estagnação a partir do algoritmo treinado com os convênios celebrados até o ano de 2013. A Figura 20, gerada a partir dos dados da Tabela 10, evidenciam uma melhoria tímida na métrica inexatidão (ϵ) dos algoritmos treinados a partir do ano de 2014, quase insignificante.

Tal fato permite responder a hipótese 4, na medida em que existe um ponto de saturação e que foi possível estabelecer o tamanho da sequência de treinamento para aprendizagem de máquina em que o incremento da volumetria não é mais significativo para a melhoria do desempenho. Por sua vez, a ocorrência de ponto de saturação no aprendizado de máquina supervisionado é corroborada por Alpaydin (2020) e Breiman (2001), Zhang(2019a). Adicionalmente, os algoritmos treinados com a volumetria saturada, ou seja, a partir de 2013 em diante, estabelecem um limiar ótimo de taxas de falsos positivos a partir do score 0,8. A Figura 20 permite afirmar que as curvas desses algoritmos possuem um ponto de inflexão precisamente a partir do *score* 0,8, uma vez que a probabilidade de ocorrência de falsos positivos permanece abaixo de 10% até esse limiar, aumentando drasticamente após esse limite.

6. CONSIDERAÇÕES FINAIS

O ciclo de vida do repasse de recursos discricionários se encerra com sua prestação de contas e conseqüente análise pelo órgão repassador, o qual opina pela aprovação ou rejeição das contas. A análise de prestação de contas é um processo moroso e exorta o emprego de recursos para sua realização, além de servidores públicos capacitados (Meireles 2019; Ferreira e Bugarin 2008). Por sua vez, o sistema “Malha Fina de Convênios” apresenta uma alternativa célere e racional para a análise de prestação de contas, configurando-se em inovação.

O objetivo desta pesquisa foi, a partir da comparação entre resultados gerados por algoritmos de Inteligência Artificial e os resultados das prestações de contas de convênios, identificar de que maneira a identidade do órgão repassador de recursos e o tamanho das sequências de treinamento são determinantes na precisão de classificação do algoritmo de aprendizado de máquina. O resultado da análise da prestação de contas de um convênio poderá ser classificada como aprovada ou rejeitada. Destarte, foi possível esclarecer as circunstâncias em que a adoção do “*Malha Fina de Convênios*” é mais assertiva por meio dos constructos *Identidade do repassador de recurso* (H1 e H2) e *Tamanho de treinamento de dados* (H3 e H4), afastando objeções contra sua implantação e validando o método dessa abordagem inovadora.

No aprendizado de máquina supervisionado *Random Forest*, os algoritmos precisam ser treinados primeiramente no conjunto de dados de treinamento e depois testados no conjunto de dados de teste. A regra geral na divisão entre conjunto de treinamento e teste é escolher cerca de 75% da amostra como dados de treinamento e o restante como dados de teste (Alpaydin, 2020). Todavia, isso não é uma regra consagrada, uma vez que a garantia de desempenho do aprendizado de máquina depende da disponibilidade de rotulação dos dados (Issa, Sun e Vasarhelyi 2016; Sun 2019) e o balanceamento dos rótulos dentro da sequência de treinamento (Bao et al. 2020; Zhang 2019a). Diante desses preceitos, os testes desta pesquisa foram realizados em duas partes, contemplando as propriedades intrínsecas de cada órgão repassador e a fidedignidade da aprendizagem conforme a agregação de dados na sequência de treinamento.

Primeiramente, investigou-se o desempenho da precisão dos algoritmos por meio da parametrização de 10 sequências de treinamento conforme a identidade do órgão concedente. Demonstrou-se, empiricamente, que, sob certas circunstâncias, alguns algoritmos treinados exclusivamente com convênios de um determinado órgão concedente possuem desempenho

melhor do que o algoritmo treinado com a população completa dos convênios (Bao et al. 2020; Cecchini et al. 2010; Tiwari e Hooda 2018). Por sua vez, essas circunstâncias se materializam sob a métrica a ser utilizada para a avaliação do desempenho, AUC ou inexatidão (ϵ). Enquanto a métrica AUC aponta 4 algoritmos com desempenho superior em relação àquele treinado com a população total de convênios, a métrica inexatidão (ϵ) não indica nenhum algoritmo treinado apenas com convênios de um determinado órgão com superioridade de desempenho em relação ao *benchmarking*. Contudo, o desempenho superior dos 4 algoritmos sob o ponto de vista da métrica AUC apresenta-se de maneira residual, quase insignificante, e sob o ônus de resultar em 6 algoritmos com desempenho inferior na perspectiva da métrica inexatidão (ϵ), sendo 2 desses muito insatisfatórios. Logo, os achados sugerem que a adoção da estratégia de treinar algoritmos separadamente conforme o órgão repassador de recursos não é vantajosa.

Em seguida, foi investigado o comportamento da precisão dos algoritmos mediante o incremento do tamanho da sequência de treinamento para a aprendizagem de máquina. Dessa forma, avaliou-se a eventual ocorrência de discrepância nos 9 algoritmos distintos gerados. Constatou-se que à medida que o tamanho da sequência de treinamento cresce, a precisão do algoritmo produzido aumenta. Ademais, o diagnóstico dos resultados obtidos indica que a adoção do modelo preditivo para estimar os resultados de um convênio não será problemática caso o tamanho da sequência de treinamento seja maior ou igual ao ponto de saturação (Breiman 2001; Domingos 2012). Nesta pesquisa, observaram-se dois pontos de saturação ligeiramente diferentes para as métricas AUC e inexatidão (ϵ). De um lado, a métrica AUC indica que a saturação ocorre com uma sequência de treinamento com convênios celebrados até o ano de 2012, consistindo em 72,18% do total (Figura 19 e Tabela 7), enquanto a métrica inexatidão (ϵ) indica a sequência composta até o ano de 2013, consistindo em 84,77% do total (Figura 20 e Tabela 7).

Com efeito, os achados dessa pesquisa validam a atual metodologia do “*Malha Fina de Convênios*”, visto que contém resultados satisfatórios baseados em testes empíricos do modelo preditivo. O Malha Fina apresenta-se como uma alternativa célere e racional para a análise de prestação de contas. Para Barzelay (1997), Bresser-Pereira (1998, 2008), Fukuyama (2004) e Power (1997, 2003a, 2003b), as tomadas de decisão em uma organização burocrática devem ser pautadas pelo custo de transação dos processos. Nessa linha, esse sistema racionaliza a utilização da força de trabalho na análise de prestações de contas pelos órgãos concedentes

mediante a adoção de um limiar de apetite a risco em que os prováveis convênios com contas rejeitáveis estariam inadvertidamente aprovados (falsos positivos). O presente trabalho mostra uma inflexão nesse limiar de risco em 0,8 (Figura 16).

Outrossim, futuras pesquisas podem avançar na aplicabilidade da metodologia do “*Malha Fina de Convênios*” em duas vertentes. A primeira vertente consiste em utilizar outras métricas, além da AUC e inexatidão (ϵ), para avaliar o desempenho do algoritmo de Inteligência Artificial. A segunda, consiste em validar a implementação da metodologia do “*Malha Fina de Convênios*” em outras modalidades de transferências de recursos da União para Entes subnacionais que não são operacionalizadas na Plataforma +Brasil. Esse campo de pesquisa mostra-se fértil, na medida em que as transferências voluntárias são quantitativamente inferiores ao volume financeiro que envolve as transferências obrigatórias (Amorim Neto e Simonassi 2013). Dados extraídos do SIAFI em maio de 2020 demonstram que, do total de recursos transferidos pela União, apenas cerca de 3% correspondem a transferências voluntárias.

Oportunamente, as pesquisas futuras poderão tirar proveito das limitações enfrentadas nesta pesquisa, adotando uma estratégia para lidar com a baixa disponibilidade de dados rotulados em classes de forma balanceada (Alpaydin 2020; Sun 2019), no caso em tela, prestações de contas de convênios aprovadas ou rejeitadas. A outra grande limitação foi a configuração de infraestrutura com grande capacidade computacional para executar as simulações.

Como qualquer disciplina, há muita "sabedoria popular" em torno da Inteligência Artificial e como seu uso pode gerar benefícios práticos. Esta pesquisa demonstrou que é possível extrair resultados dentro do arcabouço weberiano racional-legal da Administração Pública brasileira por meio de ferramentas da Ciência da Computação (Newell e Simon 1976). Isso remonta a uma situação renascentista que combina tecnologia, ciência cognitiva e necessidade humana para produzir algo que o mundo não sabia que estava faltando, constituindo um novo campo de conhecimento.

As transferências voluntárias da união não figuram como uma política pública isolada, muito menos um fim em si mesmas, haja vista que sua finalidade primordial é possibilitar, a título de cooperação, o financiamento dos serviços públicos nos entes subnacionais. Torna-se imprescindível que os critérios de equidade sobreponham-se aos objetivos políticos nos repasses decorrentes de transferências voluntárias, pois assim esse recurso contribuiria para a

diminuição das desigualdades inter-regionais na federação brasileira, conforme apontam Abrucio (2005), Arretche (2010) e Fajardo (2016). Destarte, o controle e a fiscalização dos recursos repassados precisam ser tempestivos e ágeis, concomitantes à execução da política pública, de forma a garantir a equidade e o *accountability* (Schedler 1999).

O “*Malha Fina de Convênios*”, apresenta-se como um instrumento poderoso para modificar o paradigma em que os entes subnacionais com alinhamento político com o mandatário da União, ou com maior representatividade no Congresso Nacional, recebam mais recursos, indiferentemente de outros critérios (Amorim Neto e Simonassi 2013; Brollo e Nannicini 2012; Ferreira e Bugarin 2008; Limongi e Figueiredo 2005; Meireles 2019; Soares e Melo 2016).

Em linha com os resultados já apresentados no setor privado, a difusão de métodos aplicados ao controle e mitigação de riscos na Administração Pública têm grande chance de impactar positivamente a gestão (Al-Qudah, Baniahmad e Al-Fawaerah 2013; Huang e Vasarhelyi 2019). A digitalização de serviços públicos, com o uso intensivo de Inteligência Artificial poderá se tornar prática corrente na Administração Pública como meio para enfrentar as crescentes demandas de ampliação da eficiência, oportunizando que servidores qualificados possam lidar com tarefas mais sofisticadas (Kim, Mannino e Nieschwietz 2009). Ressalta-se que os resultados desta pesquisa poderão ser incorporados no processo de trabalho dos repasses voluntários da União aos entes subnacionais, ensejando, inclusive, alterações normativas que regram o processo das transferências voluntárias da União, além de afastar o medo na adoção do “*Malha Fina de Convênios*”. Neste sentido reside a principal colaboração deste trabalho.

REFERÊNCIAS

Abrucio, Fernando Luiz. 2005. “Reforma do Estado no federalismo brasileiro: a situação das administrações públicas estaduais”. *Revista de Administração Pública*, v. 39, n.2, (2005): p. 401-420.

Abrucio, Fernando Luiz, e Cibele Franzese. “Federalismo e políticas públicas: o impacto das relações intergovernamentais no Brasil”. In *Tópicos da Economia Paulista para Gestores Públicos*, editado por Araújo e Beira, v. 1, p. 13-31. São Paulo: Fundap, 2007.

ALPAYDIN, Ethem. *Introduction to machine learning*. MIT press, 2020.

Afonso, José Roberto, Eliane Cristina Araújo, e Bernardo Guelber Fajardo. “The Role of Fiscal and Monetary Policies in the Brazilian Economy: Understanding Recent Institutional Reforms and Economic Changes.” *Quarterly Review of Economics and Finance* (2016): 41–55. <https://doi.org/10.1016/j.qref.2016.07.005>.

Al-qudah, Anas Ali, Ahmad Yahya Baniahmad, e Nawwaf Al-fawaerah. “The Impact of Information Technology on the Auditing Profession.” *Management and Administrative Sciences Review* (2013) 430: 423.

Alves, Maria Fernanda Colaço Alves, e Paulo Carlos Du Pin Calmon. “Múltiplas Chibatas? – Governança Da Política de Controle Da Gestão Pública Federal.” *ENAPG - Encontro de Administração Pública e Governança*, (2008): 1–16.

Amorim Neto, Octavio. “Gabinetes presidenciais, ciclos eleitorais e disciplina legislativa no Brasil”. *Dados*, v. 43, n. 3, (2000): p. 479-519.

Amorim Neto, Octavio, e Andrei Gomes Simonassi. “Bases Políticas Das Transferências Intergovernamentais No Brasil (1985-2004).” *Revista de Economia Política* (2013): 704–25.

Arretche, Marta. “Federalismo e Igualdade Territorial: Uma Contradição em Termos?” *Dados*

(2010): 587–620. <https://doi.org/10.1590/S0011-52582010000300003>.

Bao, Yang, Bin Ke, Bin Li, Y. Julia Yu, e Jie Zhang. 2020. “Detecting Accounting Fraud in Publicly Traded U.S. Firms Using a Machine Learning Approach.” *Journal of Accounting Research*, (2020): 199–235. <https://doi.org/10.1111/1475-679X.12292>.

Barzelay, Michael. “Central Audit Institutions and Performance Auditing: A Comparative Analysis of Organizational Strategies in the OECD”. *Governance: An International Journal of Policy, Administration, and Institutions*, (1997): 235-260.

Barzelay, Michael. *Public Management as a Design-Oriented Professional Discipline*. Cheltenham, Reino Unido. Edward Elgar Publishing Limited The Lypiatts. 2019.

Barzelay, Michael. 2018. “The Study of Public Management : Conceptualizing a Design-Oriented Social Science.” *Revista Do Serviço Público* (2018): 33–56.

Barzelay, Michael, e Raquel Gallego. “From ‘New Institutionalism’ to ‘Institutional Processualism’: Advancing Knowledge about Public Management Policy Change.” *Governance* (2006): 531–57. <https://doi.org/10.1111/j.1468-0491.2006.00339.x>.

Brasil. Controladoria-Geral da União, Secretaria Federal de Controle Interno. *Instrução Normativa nº 03, de 09.06.2017*. Aprova o Referencial Técnico da Atividade de Auditoria Interna Governamental do Poder Executivo Federal. 2017a. Disponível em https://www.cgu.gov.br/sobre/legislacao/arquivos/instrucoes-normativas/in_cgu_03_2017.pdf. Acesso em 25 mai. 2020.

Brasil. Controladoria-Geral da União, Secretaria Federal de Controle Interno. *Portaria nº 378, de 07.02.2017*. Constitui Grupo de Trabalho para avaliar as transferências voluntárias da união. 2017b.

Brasil. Controladoria-Geral da União, Secretaria Federal de Controle Interno. *Relatório de Avaliação da Gestão das Transferências Voluntárias da União*. 2018a. Disponível em

<https://auditoria.cgu.gov.br/download/11014.pdf>. Acesso em 25 mai. 2020.

Brasil. Controladoria-Geral da União e Ministério da Economia. *Instrução Normativa Interministerial nº 1, de 14 de fevereiro de 2019*. Estabelece diretrizes e parâmetros para a análise de prestação de contas simplificada. 2019a. Disponível em <http://portal.convenios.gov.br/analise-informatizada>. Acesso em 25 mai. 2020.

Brasil. Controladoria-Geral da União e Ministério do Planejamento, Desenvolvimento e Gestão. *Instrução Normativa Interministerial nº 5, de 6 de novembro de 2018*. Estabelece diretrizes e parâmetros para a análise de prestação de contas simplificada. 2018b. Disponível em <http://portal.convenios.gov.br/analise-informatizada>. Acesso em 25 mai. 2020.

Brasil. Ministério da Economia. *Plataforma +Brasil*. 2019b. Disponível em <http://plataformamaisbrasil.gov.br/>. Acesso em 25 mai. 2020.

Brasil. Tribunal de Contas da União. *Transferências Governamentais Constitucionais*. 2008. Disponível em: <<https://portal.tcu.gov.br/>>. Acesso em: 25 mai. 2020.

Breiman, Leo.. “Random Forests.” *Machine Learning*, (2001): 1–122. <https://doi.org/10.1201/9780367816377-11>.

Bresser Pereira, Luiz Carlos. “Da administração burocrática à gerencial”. *Revista do Serviço Público*, v. 47, n. 1, (1996).

Bresser Pereira, Luiz Carlos. “A reforma do estado dos anos 90: lógica e mecanismos de controle”. *Lua Nova: Revista de Cultura e Política*, (1998): 49-95.

Bresser-Pereira, Luiz Carlos. 2008. “O Modelo Estrutural de Gerência Pública.” *Revista de Administração Pública*, (2008): 391–410. <https://doi.org/10.1590/s0034-76122008000200009>.

Brollo, Fernanda, e Tommaso Nannicini. “Tying Your Enemy ’ s Hands in Close Races : The

Politics of Federal Transfers in Brazil Tying Your Enemy ' s Hands in Close Races : The Politics of Federal Transfers in Brazil”. *American Political Science Review*, (2012): 742–761. <https://doi.org/10.1017/S0003055412000433>.

Brown-Liburd, Helen, e Miklos A. Vasarhelyi. “Big Data and Audit Evidence.” *Journal of Emerging Technologies in Accounting* (2015): 1–16. <https://doi.org/10.2308/jeta-10468>.

Campana, Priscilla de Souza Pestana. “A Cultura Do Medo Na Administração Pública e a Ineficiência Gerada Pelo Atual Sistema de Controle.” *Revista de Direito*, (2017): 189–216.

Cecchini, Mark, Haldun Aytug, Gary J. Koehler, e Praveen Pathak. “Detecting Management Fraud in Public Companies.” *Management Science*, (2010): 1146–1160. <https://doi.org/10.1287/mnsc.1100.1174>.

Committee Of Sponsoring Organizations Of The Treadway Commission – COSO. *Controle Interno – Estrutura Integrada. Sumário Executivo*. Traduzido por PwC. Instituto dos Auditores Internos do Brasil, 2013.

Creswell, John W. “Projeto de Pesquisa: métodos qualitativo, quantitativo e Misto”. Porto Alegre: Artmed, 2007.

Dallaverde, Alexsandra Katia. “As transferências voluntárias no modelo constitucional brasileiro”. São Paulo: Blucher, 2016.

Domingos, Pedro.. “A Few Useful Things to Know About Machine Learning.” *Communications of the ACM*, (2012): 9–48. <https://doi.org/10.1007/978-3-642-35289-8-3>.

Dunleavy, Patrick, Helen Margetts, Simon Bastow, e Jane Tinkler. 2006. “New Public Management Is Dead - Long Live Digital-Era Governance.” *Journal of Public Administration Research and Theory*, (2006): 467–94. <https://doi.org/10.1093/jopart/mui057>.

Fajardo, Bernardo Guelber. “Vieses orçamentários em entes subnacionais: uma análise sob a ótica da estimação das receitas estaduais.” Tese de Doutorado em Administração Pública,

Fundação Getúlio Vargas, Rio de Janeiro, Brasil, 2016.

Fayol, Henry. “General and industrial management”. London, Pitman, 1949.

Fawcett, Tom.. “An Introduction to ROC Analysis.” *Pattern Recognition Letters*, (2006): 861–874. <https://doi.org/10.1016/j.patrec.2005.10.010>.

Ferreira, Ivan F. S., e Mauricio S. Bugarin. “Transferências Voluntárias e Ciclo Político-Orçamentário No Federalismo Fiscal Brasileiro.” *Revista Brasileira de Economia*, (2018): 271–300. <https://doi.org/10.1590/s0034-71402007000300001>.

Fukuyama, Francis. “A Construção de Estados. Governo e Organização no Século XXI”. Rio de Janeiro, Rocco, 2004.

Grover, Dhruv, Sebastian Bauhoff, e Jed Friedman. “Using Supervised Learning to Select Audit Targets in Performance-Based Financing in Health: An Example from Zambia.” *PLoS ONE*, (2019): 1–13. <https://doi.org/10.1371/journal.pone.0211262>.

Heber, Florence. “O Diálogo Entre Burocracia e Inovação.” *ANPAD*, (2014) 1–10.

Hood, Christopher, Oliver James e Colin Scott. “Regulation of government: Has it increased, is it increasing, should it be diminished?” *Public Administration*, v. 78, n. 2, (2000): p. 283–304.

Hooda, Nishtha, Seema Bawa, e Prashant Singh Rana. “Optimizing Fraudulent Firm Prediction Using Ensemble Machine Learning: A Case Study of an External Audit.” *Applied Artificial Intelligence*, (2020): p. 20–30. <https://doi.org/10.1080/08839514.2019.1680182>.

Huang, Feiqi, e Miklos A. Vasarhelyi. “Applying Robotic Process Automation (RPA) in Auditing: A Framework.” *International Journal of Accounting Information Systems* (2019): 100433. <https://doi.org/10.1016/j.accinf.2019.100433>.

Instituto Brasileiro de Geografia Estatística – IBGE – (2020). *Código dos Municípios IBGE*.

Disponível em <https://www.ibge.gov.br/explica/codigos-dos-municipios.php>. Acesso em: 12 abr. 2020

Issa, Hussein, Ting Sun, e Miklos A. Vasarhelyi. “Research Ideas for Artificial Intelligence in Auditing: The Formalization of Audit and Workforce Supplementation.” *Journal of Emerging Technologies in Accounting*, (2016): 1–20. <https://doi.org/10.2308/jeta-10511>.

Jackson, Bill. “Designing Projects and Project Evaluations Using the Logical Framework Approach.” *Monitoring and Evaluation Initiative: IUCN*, (1997): 1–11.

Kim, Hyo Jeong, Michael Mannino, e Robert J. Nieschwietz. “Information Technology Acceptance in the Internal Audit Profession: Impact of Technology Features and Complexity.” *International Journal of Accounting Information Systems*, (2009): 214–28. <https://doi.org/10.1016/j.accinf.2009.09.001>.

Kohavi, Ron. “A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection.” *International Joint Conference on Artificial Intelligence*, (1995): 1137–1145. <https://doi.org/10.1067/mod.2000.109032>.

Larcker, David F, e Anastasia A. Zakolyukina. “Detecting Deceptive Discussions in Conference Calls.” *Journal of Accounting Research*, (2012): 495–540.

Limongi, Fernando, e Argelina Figueiredo. “Processo Orçamentário e Comportamento Legislativo: Emendas Individuais, Apoio ao Executivo e Programas de Governo.” *Dados*, (2005): 737–76. <https://doi.org/10.1590/s0011-52582005000400002>.

Loureiro, Maria, e Fernando Abrucio. 2004. “Política e Reformas Fiscais No Brasil Recente.” *Brazilian Journal of Political Economy*, v. 24, (2004): 50–72.

Maciejewski, Mariusz. “To do more, better, faster and more cheaply: using big data in public administration”. *International Review of Administrative Sciences*, (2017): 120–135.

Margetts, Helen, e Patrick Dunleavy. “Better Public Services through E-Government: Academic Article in Support of Better Public Services through e-Government”. *Report by the Comptroller and Auditor General HC 704-III Session*, (2002). April: 1–21.

Meireles, Fernando. “Alinhamento Partidário e Demanda Por Transferências Federais No Brasil.” *Revista de Administração Pública*, (2019): 173–94. <https://doi.org/10.1590/0034-761220170282>.

Mintzberg, Henry. “*The structuring of organizations: A synthesis of the research.*” Englewood Cliffs, N.J: Prentice-Hall, 1979

Monteiro, Renato Pereira. “Análise do sistema de controle interno no Brasil: objetivos, importância e barreiras para sua implantação”. *Revista Contemporânea de Contabilidade*, (2015): 159-181.

Newell, Allen, e Hebert Simon. “Computer science as empirical inquiry: symbols and search.” *Communications of the ACM* . (1976): 113–126.

Oliveira, Antonio. “Burocratas da linha de frente: executores e fazedores das políticas públicas”. *Revista de Administração Pública*, (2012): 1551 - 1573.

Olivieri, Cecília. “Política e burocracia no Brasil: o controle sobre a execução das políticas públicas”. Tese de Doutorado em Administração Pública e Governo, Fundação Getúlio Vargas, São Paulo, Brasil, 2008.

Olivieri, Cecília. “Os controles políticos sobre a burocracia”. *Revista de Administração Pública*, (2011): 1395 – 1424.

Parker, Lee D., Kerry Jacobs, e Jana Schmitz. “New Public Management and the Rise of Public Sector Performance Audit.” *Accounting, Auditing & Accountability Journal*, (2018): 280–306. <https://doi.org/10.1108/aaaj-06-2017-2964>.

Power, Michael.. “Evaluating the Audit Explosion.” *Law and Policy*, 25, (2003a): 185–202.
<https://doi.org/10.1111/j.1467-9930.2003.00147.x>.

Power, Michael . “The Audit Society - Second Thoughts.” *International Journal of Auditing*,
4, (2003b): 111–119. <https://doi.org/10.1111/1099-1123.00306>.

Power, Michael. “The Risk Management of Nothing.” *Accounting, Organizations and Society*,
(2009): 849–855. <https://doi.org/10.1016/j.aos.2009.06.001>.

Ramos, Alberto Guerreiro. 1983. “Administração e contexto brasileiro: esboço de uma teoria
geral da administração”. Rio de Janeiro: FGV, 1983.

Schedler, Andreas. “Conceptualizing Accountability”. In *The Self-Restraining State: Power
and Accountability in New Democracies*, edited by Diamond, L. J.; Plattner, M. F. Boulder
Colorado: Lynne Rienner Publishers, 1999.

Soares, Márcia Miranda, e Bruno Guimarães De Melo. “Condicionantes Políticos e Técnicos
Das Transferências Voluntárias Da União Aos Municípios Brasileiros”. *Revista de
Administração Pública*, (2016): 539–61.

Stanisic, Nemanja, Tijana Radojevic, and Nenad Stanic. “Predicting the Type of Auditor
Opinion: Statistics, Machine Learning, or a Combination of the Two?” *The European Journal
of Applied Economics*, (2019): 1–58. <https://doi.org/10.5937/ejae16-21832>.

Sun, Ting Sophia. “Applying Deep Learning to Audit Procedures: An Illustrative
Framework.” *Accounting Horizons* (2019): 89–109. <https://doi.org/10.2308/acch-52455>.

Tenorio, Fernando Guilherme. “Weber e a burocracia.” *Revista de Serviço Público*, vol. 109,
no 4, (1981): p. 79-88.

Tiwari, Arpit.; Hooda, Nishtha. “Machine Learning Framework for Audit Fraud Data
Prediction”. *ResearchGate*, v. 7, n. 6, (2018): 164–167.

Taylor, Frederick Winslow. “*The principles of scientific management*”. New York: Harper & Brothers, 1911.

The Institute of Internal Auditors – IIA (2012). *International Standards for The Professional Practice of Internal Auditing (Standards)*. Disponível em de <https://na.theiia.org/standards-guidance/Public%20Documents/IPPF%202013%20English.pdf>. Acesso em: 10 jan. 2020.

Trevisan, Andrei Pittol, e Hans Michael Van Bellen. “Avaliação de políticas públicas: uma revisão teórica de um campo em construção”. *Revista de Administração Pública*, v. 42, n. 3, (2008), p. 529–550.

United States Government Accountability Office – GAO (2011). *Government Auditing Standards*. Disponível em <http://gao.gov/assets/590/587281.pdf>. Acesso em 25 abr. 2020.

Vasarhelyi, Miklos A., e Fern B. Halper. “The Continuous Audit of Online Systems.” *Auditing: A Journal of Practice and Theory*. (1991): 110-125.

Vasarhelyi, Miklos A., Alexander Kogan, e Brad Tuttle. “Big data in accounting: An overview”. *Accounting Horizons*, v. 29, n. 2, (2015): p. 381–396.

Wirtz, Bernd W., Jan C. Weyerer, e Carolin Geyer. 2019. “Artificial Intelligence and the Public Sector—Applications and Challenges.” *International Journal of Public Administration* (2019): 596–615. <https://doi.org/10.1080/01900692.2018.1498103>.

Zhang, Abigail Chanyuan. “Predict Audit Quality Using Machine Learning Algorithms.” *SSRN Electronic Journal*, (2019): 1–40. <https://doi.org/10.2139/ssrn.3449848>.

Zhang, Abigail Chanyuan. “Intelligent Process Automation in Audit.” *Journal of Emerging Technologies in Accounting* (2019): 69–88. <https://doi.org/10.2308/jeta-52653>.

ANEXO I - CÓDIGO PHYTON UTILIZADO NAS SIMULAÇÕES

Código para treinamento do Algoritmo de Machine Learning:

```
import pyodbc
from math import floor
from sklearn.tree import DecisionTreeClassifier
import numpy as np
import functools
from sklearn.model_selection import train_test_split
from sklearn.metrics import confusion_matrix
from sklearn.metrics import f1_score
from sklearn.metrics import recall_score
import matplotlib.mlab as mlab
import matplotlib.pyplot as plt
import pandas as pd
from sklearn.ensemble import ExtraTreesClassifier
from sklearn.model_selection import cross_validate
from sklearn.model_selection import cross_val_score
from sklearn.model_selection import cross_val_predict
import pickle
import importlib
# Connection to database to retrieve the covenants
conn = pyodbc.connect(
    r'DRIVER={ODBC Driver 13 for SQL Server};'
    r'SERVER=sdh-die-bd.cgu.local;'
    r'DATABASE=mdl_prestacao_siconv;'
    r'Trusted_Connection=yes;'
    #r'UID=;'
    #r'PWD=;'
)
query = """
    SEQUENCIA DE TREINAMENTO BANCO DE DADOS PLATAFORMA +BRASIL
    PARAMETRIZADA CONFORME O TAMANHO DA SEQUENCIA DE TREINAMENTO E A
    IDENTIDADE DO ORGAO CONCEDENTE
    """
data = pd.read_sql(query, conn)
conn.close()
columns = [column for column in data.columns if column not in ['classes_num', 'num_convenio']]
labels = data["classes_num"].values
features = data[list(columns)].values
sample_weight = np.random.RandomState(1234).rand(labels.shape[0])
```

```
f_train, f_test, c_train, c_test, sw_train, sw_test = train_test_split(features, labels, sample_weight, test_size=0.40,
random_state=1234)
#Recurso Computacional - Esforço
forest = ExtraTreesClassifier(n_estimators=1000, max_depth=None, min_samples_split=0.02, random_state=1234,
n_jobs=3, class_weight={0: 1, 1: 1000}, max_features='log2')
#A variavel et_pred é o modelo criado a partir da sequencia de treinamento da query anterior
et_pred = forest.fit(f_train, c_train, sample_weight=sw_train)
----
#Matrizes de Confusao para geracao das notas de predicao --> risco
# Matris 2 X 2 -> linha sequencia de treinamento /coluna nota atribuida pelo sistema
predictions = et_pred.predict(f_test)
print(f1_score(c_test, predictions, average=None))
confusion_matrix(c_test, predictions, labels=[0,1])
predictions_prob = et_pred.predict_proba(f_test)

# Connection to database to retrieve the covenants
conn = pyodbc.connect(
    r'DRIVER={ODBC Driver 13 for SQL Server};'
    r'SERVER=sdh-die-bd.cgu.local;'
    r'DATABASE=mdl_prestacao_siconv;'
    r'Trusted_Connection=yes;'
    #r'UID=Trusted_Connection;'
    #r'PWD=Trusted_Connection;'
)
query = """
    SEQUENCIA DE TREINAMENTO BANCO DE DADOS PLATAFORMA +BRASIL
    """
data2 = pd.read_sql(query, conn)
conn.close()
----
columns2 = [column for column in data.columns if column not in ['classes_num', 'num_convenio']]
labels2 = data2["classes_num"].values
features2 = data2[list(columns)].values
----
#Matrizes de Confusao para geracao das notas de predicao --> risco
# A Variavel et_pred eh aquela que contem o modelo criado a partir do bloco da primeira query
predictions2 = et_pred.predict(features2)
print(f1_score(labels2, predictions2, average=None))
confusion_matrix(labels2, predictions2, labels=[0,1])
----
#Imprimindo as variaveis do teste feito no modelo
predictions_prob2 = et_pred.predict_proba(features2)
```




Revista Cadernos de Finanças Públicas, Brasília, v. 02, n. 1, p. 1-81,
Edição Especial 2021

```
data2['scores0']=predictions_prob2[:,0]
```

```
data2['scores1']=predictions_prob2[:,1]
```

```
data2.to_csv('Modelo_Todos_Orgaos_resultado_para_todos_orgaos.csv')
```