

Revista

Cadernos de Finanças Públicas

03 | 2024



TESOURONACIONAL

PREVISÃO DE FRAUDE EM LICITAÇÕES NO BRASIL

Daniel Vitor Tartari Garruti

Universidade Federal de Uberlândia (UFU)

Vinicius Moraes

Universidade Federal de Uberlândia (UFU)

Pablo Rogers

Universidade Federal de Uberlândia (UFU)

Flavio Barboza

Universidade Federal de Uberlândia (UFU)

RESUMO

Os portais eletrônicos de licitação modernizaram a aquisição de bens e serviços, tornando-a mais eficiente, competitiva e transparente. Para acompanhar essas mudanças, a fiscalização também deve se adaptar. No contexto acadêmico, observa-se o avanço do uso de técnicas de Inteligência Artificial, como o *Random Forest*, para prever fraudes em licitações. Este estudo aplicou o modelo em dados de Licitações de 2022 para prever a incidência de multas em empresas. O modelo demonstrou eficiência, com F1 Score médio mensal de 78,5% e anual de 80%. O recall foi de 90% para ambos os períodos. Em Janeiro, os melhores resultados foram obtidos (*recall* = 1,00 e *F1 Score* = 0,96), enquanto em Novembro houve queda significativa (*recall* = 0,47 e *F1 Score* = 0,54). Destaca-se o impacto potencial deste estudo não apenas na academia, mas também para cidadãos e gestores públicos, oferecendo uma ferramenta eficaz na detecção de potenciais desvios de conduta.

Palavras-Chave: Licitações, Fraudes, Serviço Público, Inteligência Artificial, *Random Forest*.

JEL: C53, D73, E66.

SUMÁRIO

1. INTRODUÇÃO	4
2. ESTUDOS RELACIONADOS	6
3. METODOLOGIA	9
3.1. <i>Coleta de dados</i>	9
3.2. <i>Random Forest</i>	14
3.3. <i>Algoritmo</i>	15
3.4. <i>Análises dos resultados</i>	16
4. RESULTADOS E DISCUSSÕES	18
5. CONCLUSÃO	25
REFERÊNCIAS BIBLIOGRÁFICAS	28

1. INTRODUÇÃO

A atividade econômica do Estado desempenha um papel crucial para impulsionar a demanda efetiva, estimulando a produção e criação de empregos no país. No Brasil, as contratações realizadas pelo setor público entre 2006 e 2017, considerando todos os entes federativos, foram em média de 12,5% do PIB. Deste total, 6,8% correspondem a compras da União, das quais a Petrobras sozinha respondeu por 4,2% (RIBEIRO; INÁCIO, 2019).

As contratações do Estado devem ser realizadas por meio de licitação (BRASIL, 1988). Nesse sentido, o sistema de leis que regulamenta as contratações públicas atualmente define o poder de compra estatal como um instrumento capaz de efetivar políticas públicas para redução das desigualdades, promoção do desenvolvimento socioeconômico, inovação e sustentabilidade (REJEB et al., 2023; ZAGO, 2018). A licitação é um procedimento formal da Administração Pública, regulamentado pela Lei nº 14.133/2021, no qual empresas são convocadas a apresentar propostas para a oferta de bens e serviços, conforme estabelecido em edital (BRASIL, 2021).

As aquisições públicas são um dos aspectos do Estado em que a probabilidade de ocorrer corrupção, fraude e conluio é maior (MAVIDIS; FOLINAS, 2022; GUNASEGARAN; BASIRUDDIN; RIZAL, 2023; REJEB et al., 2023). Condutas infratoras incluem atos que comprometam a isonomia e a competitividade como combinação de preços ou simulação de propostas; alterações indevidas no contrato, pagamentos em desacordo com as normas contratuais, afastamento indevido de licitante por meio de fraude ou oferecimento de vantagem, atraso injustificado na execução de obras ou serviços, inexecução do contrato, corrupção, obstrução da fiscalização. A transparência nas contratações, a capacitação da equipe responsável e a atuação do controle interno e externo são aspectos centrais na prevenção de fraudes. Na literatura, o principal método de detecção das fraudes é a utilização de *Red Flags*/Alertas para que os auditores aprofundem suas análises sobre cada caso. Outras formas de detecção estão associadas à disponibilização de canais de denúncia e à utilização de equipes multidisciplinares para investigação (GUNASEGARAN; BASIRUDDIN; RIZAL, 2023).

As normas brasileiras também estabelecem uma série de sanções administrativas e penais para garantir a transparência, moralidade e eficiência nesses processos. A Lei nº 8.666/1993, conhecida como Lei de Licitações e Contratos, prevê punições como advertência, multa, suspensão temporária de participação em licitação e declaração de inidoneidade para licitar ou contratar com a Administração Pública (BRASIL, 1993). A Lei do Pregão (Lei nº 10.520/2002) e o Regime Diferenciado de Contratações Públicas (Lei nº 12.462/2011) também estipulam

sanções semelhantes, enquanto a Nova Lei de Licitações e Contratos Administrativos (Lei nº 14.133/2021) introduz novas disposições, mantendo uma estrutura similar de sanções administrativas (BRASIL, 2002; BRASIL, 2011; BRASIL, 2021). A Lei Anticorrupção (Lei nº 12.846/2013) estabelece punições para pessoas jurídicas envolvidas em atos lesivos contra a administração pública, enquanto o Código Penal tipifica crimes como fraude em licitação, peculato e corrupção (BRASIL, 2013). A implementação dessas leis visa promover a integridade no ambiente corporativo e governamental, além de combater a corrupção no país, alinhando-se aos esforços internacionais de governança e transparência.

No entanto, a aplicação eficaz dessas leis enfrenta desafios significativos. A complexidade dos processos de licitação e a grande quantidade de transações tornam a detecção de irregularidades uma tarefa árdua. A identificação precoce de possíveis fraudes e atos corruptos é essencial para a eficácia das medidas preventivas e punitivas. Nesse contexto, a tecnologia desempenha um papel crucial ao fornecer ferramentas avançadas para monitorar e analisar os dados das contratações públicas.

A adoção de sistemas eletrônicos, trouxe melhorias na eficiência e na eficácia do setor público, em especial quanto à transparência, à governança, à melhor utilização dos recursos públicos, à promoção do desenvolvimento regional e da competição entre os fornecedores (ADOBOR; YAWSON, 2023; GADOUR, 2024; GUNASEGARAN; BASIRUDDIN; RIZAL, 2023; MAVIDIS; FOLINAS, 2022). As compras eletrônicas (*e-procurement*) se tornaram comuns em todas as esferas de governo, destacando-se o portal Comprasnet. A Administração Pública também tem adotado com cada vez mais frequência ferramentas de inteligência artificial (IA). No Supremo Tribunal Federal, destacam-se iniciativas como o VitorIA e o Projeto Victor, além do RAFA 2030 (Redes Artificiais Focadas na Agenda 2030) (SUPREMO TRIBUNAL FEDERAL, 2021; SUPREMO TRIBUNAL FEDERAL, 2023a; SUPREMO TRIBUNAL FEDERAL, 2023b). Já o Tribunal de Contas da União emprega as IAs generativas, incluindo o chatbot Zello, o ChatTCU e o Copilot TCU (TRIBUNAL DE CONTAS DA UNIÃO, 2024; SECOM TCU, 2020). No campo das compras públicas, a principal iniciativa é o software Alice (Analisador de Licitações, Contratos e Editais), desenvolvido pela Controladoria Geral da União (2024) (BIAGINI, 2024).

Esses sistemas eletrônicos associados às Inteligências Artificiais (IAs) são ferramentas fundamentais para dificultar a ocorrência de fraudes nas contratações (ADOBOR; YAWSON, 2023; GADOUR, 2024; GUNASEGARAN; BASIRUDDIN; RIZAL, 2023). Os modelos de aprendizado de máquina (ML - *machine learning*) oferecem uma abordagem inovadora para

prever condutas ilegais na aquisição pública, utilizando análise de grandes conjuntos de dados para identificar padrões e variáveis associadas a comportamentos corruptos. Essas previsões podem ser utilizadas como pontuações de risco por agências anticorrupção e partes contratantes para orientar a seleção de operações de aquisição que precisam de monitoramento mais rigoroso ou auditorias.

Diante disso, este estudo propõe um modelo de previsão de fraudes (alertas) em licitações no Brasil utilizando a técnica *Random Forest* (RF). Este estudo tem utilidade para os agentes públicos que atuam nas contratações públicas, pois permite a identificação de casos com maiores chances de irregularidades de forma preditiva, facilitando a análise em tempo real (SICILIANI et al., 2023). Tendo em vista as competências desempenhadas pelo controle externo no âmbito da governança pública e a experiência no uso de ferramentas de Inteligência Artificial (IA), este trabalho será especialmente útil no trabalho de auditoria. Sob a perspectiva de contribuições acadêmicas, esta pesquisa preenche uma lacuna na literatura relacionada às compras públicas, pois busca explorar o campo da aplicação de uma ferramenta de IA, que é um campo com poucos trabalhos, conforme identificado por Rejeb et al. (2023) e Gunasegaran, Basiruddin e Rizal (2023).

2. ESTUDOS RELACIONADOS

Nai et al. (2023) aplicaram técnicas de ML e sistemas de recomendação para melhorar a eficiência e transparência dos processos de aquisição pública, bem como para prever possíveis reclamações nos tribunais administrativos. Os resultados do estudo mostraram que é possível utilizar técnicas de ML em conjuntos de dados jurídicos reais relacionados aos processos de aquisição na Itália para prever reclamações nos tribunais administrativos com base nas características dos processos de aquisição. Os autores desenvolveram um modelo de ML capaz de reconhecer anomalias em processos de aquisição e identificar as características mais relevantes para a classificação dos processos com ou sem reclamações. O modelo apresentou uma acurácia de 80% na previsão de reclamações, o que demonstra a eficácia da abordagem proposta.

Além disso, os autores desenvolveram um sistema de recomendação para retornar processos de aquisição semelhantes e encontrar empresas para licitantes, com base nos requisitos do processo de aquisição. O sistema de recomendação foi desenvolvido utilizando técnicas de ML e redes neurais profundas, e apresentou resultados promissores na identificação de processos de aquisição semelhantes e empresas para licitantes (NAI et al., 2023).

Os resultados do estudo destacam a aplicabilidade e o potencial das técnicas de ML e sistemas de recomendação na melhoria dos processos de aquisição e na previsão de possíveis reclamações nos tribunais administrativos. Essas técnicas podem contribuir para a eficiência e transparência no setor público, permitindo que as autoridades identifiquem possíveis anomalias nos processos de aquisição e tomem medidas preventivas para evitar reclamações futuras. Além disso, o estudo demonstra a importância de explorar conjuntos de dados jurídicos abertos para melhorar os sistemas, procedimentos e serviços do setor público (NAI et al., 2023).

Oliveira et al. (2023) identificaram indícios de fraudes por meio da análise das interações entre os participantes das licitações, como empresas licitantes e seus sócios, modelando essas interações como uma rede social. Essa metodologia foi exitosa em reduzir um grande volume de dados em um indicador, que possibilitou ranquear as contratações onde o risco de fraude era maior. Dessa forma, o trabalho de auditoria dos especialistas do Ministério Público de Minas Gerais pode ser mais bem orientado para análise daqueles casos onde o risco de fraude seja maior.

Gallego et al. (2021) identificaram a capacidade demonstrada pelos modelos de ML em prever condutas ilegais em contratos de aquisição pública, fornecendo pontuações de risco que podem ser utilizadas para priorizar contratos que exigem monitoramento mais rigoroso. Além disso, o estudo identificou variáveis-chave, como tamanho e duração dos contratos, atrasos na implementação e padrões setoriais, que estão significativamente associadas à corrupção e ineficiência.

Com isso, os resultados desse estudo oferecem sugestões/propostas interessantes para a formulação de políticas direcionadas e reformas regulatórias para mitigar o risco de corrupção na aquisição pública. A pesquisa também destaca a importância da transparência e responsabilidade no processo de aquisição, propondo o uso de plataformas baseadas na web para registrar e relatar transações públicas. Ademais, a combinação de métodos de ML com técnicas tradicionais de inferência causal é apresentada como uma abordagem promissora para fortalecer as políticas de combate à corrupção e aprimorar a integridade dos processos de aquisição pública.

Decarolis e Giorgiantonio (2020) enfatizam a importância de melhorar a coleta de dados e a transparência em contratações públicas como meios eficazes de combate à corrupção. Eles recomendam a adaptação dos indicadores de *red flags* de acordo com as características específicas de cada região e setor, dada a variabilidade nos padrões de corrupção observados. Além disso, destacam o potencial do uso de ferramentas analíticas avançadas, como o ML, na administração pública para detectar e prevenir a corrupção de forma mais eficaz.

Utilizando modelos de ML, eles demonstraram que certos indicadores, como a unicidade do licitante e a falta de competição, são eficazes na previsão de corrupção. Observaram também variações significativas desses sinais entre diferentes regiões e tipos de contratos na Itália, sugerindo a necessidade de abordagens regionais específicas na detecção de corrupção (DECAROLIS; GIORGIANTONIO, 2020).

Decarolis e Giorgiantonio (2020) indicam que para melhorar os processos de aquisição e reduzir a corrupção em licitações públicas devem ser identificados os indicadores de risco, deve haver o aprimoramento da competição entre empresas e a implementação de práticas anticorrupção. Essas medidas podem ajudar a fortalecer a integridade e a transparência nas licitações públicas, contribuindo para a redução da corrupção. Além disso, foi identificada uma forte correlação entre certos sinais vermelhos e o aumento do risco de corrupção em contratações públicas.

Henrique, Sobreiro e Kimura (2020) compararam as principais técnicas de ML, incluindo RF, para classificar contratos públicos brasileiros com relação ao risco de não conformidade. Eles também analisaram regressão logística, K-vizinhos mais próximos (*K-Nearest Neighbours*, KNN), análise discriminante e máquina de vetores de suporte (*Support Vector Machine*, SVM). Empresas que receberam quaisquer sanções graves foram classificadas como de alto risco. As variáveis independentes utilizadas foram: dados de aquisições, quantidades de itens, lances em negociações eletrônicas, valor médio aprovado no processo de licitação, número de órgãos com os quais contratos foram concluídos, número de entidades federais com as quais órgãos contrataram com a empresa, número de itens vencedores no processo de licitação, desclassificações/irregularidades sofridas pelas empresas, número médio de empregados e seus salários médios, renda recebida pelos sócios, número de atividades comerciais registradas junto à SRF, idade da empresa, impostos pagos às autoridades, tamanho da empresa (MESB e NPO), soma dos valores negociados em contratos na condição favorecida, número de sócios e empregados que também eram servidores públicos ou com algum tipo de função remunerada, e quaisquer doações para campanhas políticas. Os dados foram obtidos das bases COMPRASNET e Sistema de Cadastramento Unificado de Fornecedores (SICAF), Relação Anual de Informações Sociais (RAIS) e no Tribunal Superior Eleitoral (TSE). O modelo utilizou 500 árvores de decisão.

De forma similar, Sá, Pessanha e Alves (2024) também compararam os principais algoritmos de classificação, incluindo RF, juntamente com regressão logística, KNN, redes neurais e SVM, para identificar a propensão de fraude nos processos de contratação do Governo do Estado do Rio de Janeiro. A variável dependente utilizada também foi sanção sofrida pelo fornece-

dor. As variáveis independentes foram: CNPJ do fornecedor, tamanho empresa (Microempresa ou Empresa de Pequeno Porte), quantidade e valor dos contratos registrados para o fornecedor, número de licitações em que o fornecedor participou, total de licitações que o fornecedor venceu, quantidade de compras diretas registradas para o fornecedor, valor de compras diretas registradas para o fornecedor, capital social, total de CNAEs, distância entre a sede do fornecedor e o local da realização do serviço, anos de criação da empresa, quantidade de empresas em que o fornecedor possui sócios em comum, situação na Receita (ativa ou inapta/suspensa), quantidade de contratos firmados com o fornecedor cujos números iniciais falharam no teste de Benford, média das idades dos sócios da empresa. Os dados foram extraídos do Portal da Transparência do Estado do Rio de Janeiro e do Portal da Transparência da União. O modelo foi configurado limitando-se a profundidade máxima da árvore para 8.

Ambos os trabalhos de Henrique, Sobreiro e Kimura (2020) e Sá, Pessanha e Alves (2024) obtiveram bons resultados para o RF, mas não calcularam as principais métricas de qualidade do modelo de classificação, como precisão, recall e F1 Score. De maneira similar, os trabalhos acadêmicos de Faria (2023), Lopes (2019) e Silva (2022) também compararam modelos de classificação, incluindo o RF, na detecção de fraudes em licitações na administração pública federal brasileira. Os resultados de Faria (2023) e Silva (2022) foram conflitantes com os resultados de Lopes (2019). O modelo de RF teve a melhor performance nos estudos de Faria (2023) e Silva (2022), enquanto obteve a pior performance nos estudos de Lopes (2019).

Reconhecendo o potencial da técnica para os órgãos de fiscalização brasileiros, este trabalho busca avaliar a eficácia do algoritmo Random Forest na detecção de fraudes em licitações usando os dados do Portal Transparência da Controladoria Geral da União.

3. METODOLOGIA

3.1. Coleta de dados

Para conduzir esta pesquisa, os dados foram obtidos por meio do Portal Transparência, vinculado à Controladoria-Geral da União, durante o período de 01/01/2022 a 31/12/2022. Foi necessário acessar quatro planilhas correspondentes a cada mês, disponíveis na seção de Licitações e Contratos, para a obtenção das informações, conforme preconizado por Gallego, Rivero e Martínez (2021). Os dados coletados e o local no site podem ser visualizados no Quadro 1.

Quadro 1: Dados coletados Portal Transparência da Controladoria-Geral da União.

VARIÁVEL	DESCRIÇÃO	PLANILHA	ABA
Valor Licitação	Valor monetário	Licitação	Licitações
Nome UG	Comprador		
Objeto	Processo de aquisição		
Modalidade Compra	Processo de aquisição		
Data Resultado Compra	Data de compra		
UF	Estado		
Número Licitação	Valor numérico da licitação		
Código UG	Código do comprador		
Nome Vencedor	Vencedor da licitação		
Número Licitação	Valor numérico da licitação		
Código UG	Código do comprador	Participantes Licitação	
Descrição Item Compra	Item negociado		
Número Licitação	Valor numérico da licitação		
Código UG	Código do comprador	Compras	Contratos
Data Assinatura Contrato	Data		
Data Início Vigência	Data		
Data Fim Vigência	Data		
Data Publicação DOU	Data		
Data Assinatura Contrato	Data		
Número Licitação	Valor numérico da licitação		
Código UG	Código do comprador		

Fonte: Elaboração própria

Após a obtenção dos dados mencionados, procedeu-se à combinação das variáveis "Número de Licitação" e "Código UG", resultando em uma nova variável denominada "Número Código", para a consolidação dos dados em um único quadro. Tal procedimento se fez essencial devido à identidade do número de licitação entre diversas entidades. Assim, ao agrupar o número da licitação com o código da entidade, obtivemos uma numeração única para cada processo.

Dessa forma, foram obtidas 16 variáveis independentes empregadas na pesquisa, as quais estão relacionadas no Quadro 2. Na literatura, observa-se uma ampla diversidade nas variáveis utilizadas para estudar fraudes em licitações. Contudo, a maior parte dos estudos não foca explicitamente em explicar os motivos pelos quais as empresas se tornam desqualificadas. Além disso, a seleção de variáveis independentes frequentemente não se baseia em uma lógica estrita de causalidade. Em vez disso, muitas pesquisas optam por abordagens exploratórias ou descritivas, que visam identificar padrões de comportamento sem necessariamente estabelecer relações causais diretas entre as variáveis, como no trabalho de Henrique, Sobreiro e Kimura (2020).

Nesse sentido, o trabalho de Gallego, Rivero e Martínez (2021) se destaca ao propor um

conjunto de variáveis independentes básicas, comum a vários outros estudos (ver referências Quadro 2), detalhando uma interpretação para o modelo e para a importância das variáveis. Esse conjunto inclui as variáveis (valor, entidades, empresas, localização, data, tipo de construção) apontadas por peritos criminais como informações que geralmente levam a uma boa estimativa sobre o risco de aquisição de fraude (LIMA et al., 2020). Assim, este trabalho utilizou principalmente o conjunto de variáveis preditoras proposto por Gallego, Rivero e Martínez (2021), aplicadas no contexto brasileiro. Os padrões de risco em potencial associados às variáveis são detalhados a seguir:

- **Orçamento:** Comparação entre o orçamento planejado e o valor final do contrato pode indicar sobrepreços. Valores muito acima do esperado podem sugerir conluio ou corrupção.

- **Tipo/Nível da entidade:** Entidades governamentais de diferentes níveis ou tipos podem ter diferentes padrões de risco de fraude. O nível de escrutínio público e regulatório pode variar significativamente entre diferentes níveis governamentais

- **Objetivo da licitação:** O propósito específico pode estar associado a certos riscos de fraude.

- **Tipo de processo:** Diferentes processos (pregão, concorrência, etc.) têm diferentes níveis de transparência e risco associado. Processos menos transparentes podem ser mais suscetíveis a fraudes.

- **Tipo de fornecedor:** Fornecedores com histórico de irregularidades ou aqueles recém-criados podem aumentar o risco de fraude.

- **Dia/Mês/Ano que o contrato foi assinado:** Datas próximas ao fim de mandatos políticos ou datas específicas podem ser usadas para apressar contratações sem o devido processo legal.

- **Período de execução:** Projetos com períodos de execução longos ou curtos podem indicar má gestão ou esforços para desviar fundos.

- **Período de espera:** Indicador de transparência e publicidade sinalizando a diligência e pontualidade das autoridades ao publicar informações sobre contratos. Um período de espera muito curto entre o anúncio e a decisão pode indicar falta de concorrência adequada

- **Intervalo entre a data de assinatura e a data de concessão:** Intervalos muito curtos podem sugerir que o fornecedor já estava predestinado a ganhar, independentemente das propostas concorrentes.

- **Intervalo entre a data de início da execução e a data de assinatura:** Intervalos muito curtos podem indicar preparação insuficiente, enquanto intervalos longos podem ser tentativas

de adiar a execução sem justificativas válidas.

- **Distância para a eleição mais próxima:** Contratos assinados muito próximos a eleições podem estar tentando influenciar resultados eleitorais ou aproveitar a mudança de administração para passar despercebidos.

- **Categoria do produto:** Certas categorias de produtos podem ser mais propensas a fraudes, especialmente se envolverem grandes somas de dinheiro ou se forem de natureza mais técnica e difícil de avaliar.

- **Estado:** Variações regionais em termos de governança e transparência podem afetar o risco de fraude. Estados com histórico de corrupção podem ter mais incidências de fraudes em licitações. Além disso, a distância entre a sede registrada do fornecedor e o local de realização do serviço podem indicar fraude.

Quadro 2: Variáveis organizadas por tipo, seguindo o formato padrão da gestão pública. As variáveis destacadas em negrito foram as utilizadas no modelo finais.

VARIÁVEIS INDEPENDENTES	Referências
Orçamento	Gallego, Rivero e Martínez (2021) Lima et al. (2020)
Tipo da entidade	Aldana, Falcón-Cortés e Larralde (2022) Gallego, Rivero e Martínez (2021) Lima et al. (2020)
Nível da entidade	
Objetivo da licitação	Gallego, Rivero e Martínez (2021)
Tipo de processo	Aldana, Falcón-Cortés e Larralde (2022) Gallego, Rivero e Martínez (2021)
Tipo de fornecedor	Aldana, Falcón-Cortés e Larralde (2022) Sá, Pessanha e Alves (2024)
Dia que o contrato foi assinado	Gallego, Rivero e Martínez (2021) Lima et al. (2020)
Mês que o contrato foi assinado	
Ano que o contrato foi assinado	
Período de execução	Aldana, Falcón-Cortés, Larralde (2022) Gallego, Rivero e Martínez (2021).
Período de espera	Gallego, Rivero e Martínez (2021)
Intervalo entre a data de assinatura e a data de concessão	Gallego, Rivero e Martínez (2021)
Intervalo entre a data de início da execução e a data de assinatura	Gallego, Rivero e Martínez (2021)
Distância para a eleição mais próxima	Aldana, Falcón-Cortés e Larralde (2022) Gallego, Rivero e Martínez (2021) Henrique, Sobreiro e Kimura (2020) Velasco et al. (2021)
Categoria do produto	Gallego, Rivero e Martínez, (2021) Lima et al. (2020)
Estado	Abreu, Pereira e Gomes-Jr (2024) Gallego, Rivero e Martínez (2021) Sá, Pessanha e Alves (2024)
VARIÁVEL DEPENDENTE	
Recebeu penalidade registrada na controladoria geral	Gallego, Rivero e Martínez (2021) Henrique, Sobreiro e Kimura (2020) Sá, Pessanha e Alves (2024)

Fonte: Elaboração própria.

A variável dependente foi definida como as instituições que receberam penalidade da Controladoria Geral no ano de 2022. Sua escolha foi baseada nos trabalhos de Gallego, Rivero e Martínez (2021), Henrique, Sobreiro e Kimura (2020) e Sá, Pessanha e Alves (2024). A utilização do registro de penalidade da Controladoria Geral como variável dependente para indicar fraude em licitações pode não capturar todos os casos de fraude existentes. De fato, pode haver situações de fraude que não resultam na aplicação de penalidade devido a diversos fatores, como limitações nas investigações, falta de provas conclusivas ou até mesmo atrasos no processo de fiscalização. No entanto, a escolha dessa variável dependente foi baseada na necessidade de um indicador objetivo e verificável de comportamento fraudulento, conforme adotado em estudos anteriores na literatura. Além disso, a aplicação de penalidade é um forte indicativo de violações significativas e detectadas, o que a torna uma métrica confiável para os casos em que a fraude é efetivamente identificada e penalizada. Aprimoramentos posteriores do modelo deverão incorporar outras variáveis (como investigações em andamento e denúncias) que indiquem irregularidades ampliando assim o espectro de detecção de fraudes para além das penalidades já confirmadas.

Analisando o Quadro 3 é possível analisar os dados descritivos das variáveis preditoras a serem utilizadas neste estudo.

Quadro 3: Dados descritivos das variáveis.

VARIÁVEIS	OBSERVAÇÕES	VALORES EXCLUSIVOS	VALOR MAIS COMUM	FREQUÊNCIA
Valor licitação	109196	52069	0	6195
Nome UG	109196	2005	Universidade Federal do Rio Grande do Sul	1510
Objeto	109196	98010	Informação Protegida por sigilo nos termos da...	1012
Modalidade compra	109196	10	Dispensa de Licitação	59614
Data resultado compra	109196	345	15/12/2022	886
UF	109196	28	RJ	22505
Nome vencedor	950270	54509	J. J. VITALLI	7620
Data assinatura contrato	35613	1056	01/09/2022	486
Data início vigência	35613	1111	01/09/2022	676
Data fim vigência	35300	2162	31/12/2022	1194
Data publicação DOU	35613	270	26/12/2022	286
Descrição assinatura contrato	35613	1056	01/09/2022	486
Descrição item compra	4442683	10926	Manutenção/Reforma Predial	72103

Fonte: Elaboração própria.

3.2. *Random Forest*

Observa-se que as técnicas de IA têm demonstrado avanços significativos na detecção de fraudes em licitações (DECAROLIS; GIORGIANTONIO, 2020; GALLEGO et al., 2021; NAI et al., 2023; OLIVEIRA et al., 2023). Diante desse cenário, optou-se por empregar uma dessas técnicas para a previsão no presente estudo, especificamente o *Random Forest* (RF – ou Floresta Aleatória em português).

A RF é formada por múltiplas árvores de decisão, analogamente a uma floresta. O processo de criação das árvores de decisão começa com a geração aleatória de subconjuntos de dados a partir dos originais, chamados de amostra *bootstrap*. Em seguida, a aleatoriedade é novamente empregada para selecionar as variáveis preditoras (atributos), formando os nós de

decisão. Os nós são definidos com base nas características que melhor separam os dados de acordo com o resultado desejado, estabelecendo pontos de corte nas variáveis independentes de forma a maximizar a diferenciação entre as categorias. Métricas como impureza de Gini, ganho de informação ou erro quadrático médio (MSE) podem ser utilizadas para avaliar a qualidade da divisão. A separação recursiva dos dados forma nós (majoritariamente) puros contendo dados de apenas uma classe, chamados nós folha. Esse procedimento é repetido para formar várias árvores de decisão, cada uma com estruturas de características distintas, resultando em modelos diferentes. Por fim, para obter uma predição final, o algoritmo percorre cada árvore e determina o resultado de cada variável, selecionando o resultado mais frequente através de um processo de "votação".

O desenvolvedor do modelo, Breiman (2001) destaca as vantagens da técnica RF, salientando sua precisão, capacidade de lidar com *outliers* e ruídos, e facilidade de implementação. Ele ressalta que a técnica fornece estimativas adequadas para variáveis internas, como erro, força, correlação e importância das variáveis, enquanto evita problemas de *overfitting* (falha na generalização), garantindo confiabilidade na previsão de resultados futuros.

3.3 Algoritmo

O modelo RF foi implementado em Python 3, utilizando a classe *RandomForestClassifier* do módulo *sklearn.ensemble* da biblioteca *Scikit-Learn* (SCIKIT-LEARN, 2024). Após a verificação da distribuição das classes, os dados foram balanceados para equilibrar as classes sub-representadas e sobre-representadas, reduzindo o número de exemplos na classe majoritária. Os dados foram divididos em conjuntos de 70% para treinamento e de 30% para teste. O modelo foi configurado com 100 árvores ($n_estimators=100$) e com $random_state=42$. A precisão do modelo foi avaliada por meio da matriz de confusão e do relatório de classificação. Também, foi feita a validação cruzada estratificada para verificar a eficácia do modelo em termos de sua capacidade de generalizar para novos dados, dividindo o conjunto de dados em 5 partes ($n_splits = 5$). Por fim, analisou-se quais características são mais importantes para a previsão de fraudes e a curva ROC foi gerada.

As análises foram feitas inicialmente por mês, devido a restrições computacionais (processador com 2 núcleos e RAM 4GB) para processar o volume de dados. Posteriormente, com um computador mais potente (processador com 12 núcleos e RAM 32GB) foi possível expandir a análise para o intervalo anual, permitindo uma avaliação mais abrangente.

3.4 Análises dos resultados

A performance do modelo de classificação implementado foi avaliada analisando os erros tipo I e II, acurácia, precisão, sensibilidade (*recall*), pontuação F1 (*F1 Score*), além de compará-los com os resultados dos estudos mencionados no referencial teórico.

O Quadro 4, a seguir, apresenta a contagem de eventos e previsões de forma que é possível observar a qualidade do modelo preditivo, também conhecida como matriz de confusão. A matriz de confusão permite visualizar de forma clara os diferentes tipos de erros (tipo I e tipo II) e os resultados corretos do modelo. Isso facilita a análise da precisão do modelo em prever fraudes como também identificar processos que cumprem os pressupostos da legalidade, ou seja, que não se verifica evidências de fraudes.

Quadro 4: Matriz de Confusão, a qual apresentará os resultados das previsões e a contagem de acertos e erros que são usados para avaliar a qualidade preditiva.

Real versus previsto	Previsto como Fraude	Previsto como Não Fraude
Fraude ocorreu	Verdadeiro Positivo (TP)	Falso Negativo (FN) <i>Erro Tipo II</i>
Fraude Não ocorreu	Falso Positivo (FP) <i>Erro Tipo I</i>	Verdadeiro Negativo (TN)

Fonte: Elaboração própria.

De maneira mais específica, o erro tipo I (dado pela Equação 1) ocorre quando a fraude não foi detectada no processo licitatório, mas é apontada erroneamente pelo modelo de previsão. Errar neste caso é considerado menos oneroso, uma vez que se assemelha a um sinal de alerta para verificação mais cuidadosa e, portanto, parte integrante (e quase natural) das atividades dos órgãos fiscalizadores. Já o erro do tipo II (Equação 2) ocorre quando a fraude é constada, mas não é prevista pelo modelo.

$$\text{Erro do Tipo I} = \frac{FP}{FP+TN} \quad (1)$$

$$\text{Erro do Tipo II} = \frac{FN}{FN+TP} \quad (2)$$

A acurácia (Equação 3) representa a média geral de acertos do modelo. No entanto, ela é

uma medida insuficiente, especialmente para classes desbalanceadas. Uma análise mais aprofundada da eficácia do modelo foi feita analisando também precisão e sensibilidade.

$$\text{Acurácia} = \frac{TN+TP}{TP+FP+TN+FN} \quad (3)$$

Precisão e sensibilidades (Equações 4) são métricas complementares. A escolha de qual métrica enfatizar depende do contexto específico da aplicação e do equilíbrio desejado entre detectar positivos e evitar falsos positivos. Na detecção de fraudes em licitações, a sensibilidade (*recall*) é frequentemente considerada mais crítica. Falsos negativos, ou seja, fraudes que não são detectadas, estão associadas a perdas financeiras, danos a reputação e favorece as redes de corrupção. Priorizar a sensibilidade ajuda a garantir que a maioria das fraudes seja detectada, minimizando a chance de fraudes passarem despercebidas. Embora menos crítica, a precisão ainda é importante, especialmente em contextos onde o volume de transações é alto. Altas taxas de falsos positivos podem levar a uma sobrecarga operacional, aumentando os custos com investigações desnecessárias e potencialmente prejudicando a experiência das empresas/fornecedores, caso transações legítimas sejam frequentemente bloqueadas ou atrasadas. Aumentar a precisão reduz o número de falsos positivos, o que pode ajudar a manter a eficiência operacional.

Assim, a Pontuação F1 é particularmente valiosa porque calcula a média harmônica entre precisão e sensibilidade (*recall*), permitindo a identificação do modelo que oferece o melhor custo-benefício. A pontuação F1 varia de 0 a 1, onde 1 é o melhor valor possível e 0 é o pior. Um valor de 1 indica perfeição na classificação, com precisão e sensibilidade perfeitas, o que significa que todos os positivos verdadeiros e nenhum falso positivo foram identificados. Já um valor de 0 indica que o modelo falhou completamente em identificar positivos verdadeiros, ou que todos os resultados positivos são falsos positivos.

$$\text{Precisão} = \frac{TP}{TP+FP} \quad \text{Sensibilidade} = \frac{TP}{TP+FN} \quad (4)$$

$$\text{F1_Score} = 2 \cdot \frac{\text{Precisão} \times \text{Sensibilidade}}{\text{Precisão} + \text{Sensibilidade}} \quad (5)$$

Além disso, a técnica de RF possui uma propriedade interessante que permite examinar quais as variáveis preditoras foram mais importantes para realizar as previsões. Assim, esta foi também uma parte da análise que pode ser consideravelmente útil para que seus usuários te-

nham mais atenção em determinados sinalizadores e, assim aumentar sua eficiência e aprimorar o processo de fiscalização nos processos de licitação.

A curva *Receiver Operating Characteristic* (ROC) permite avaliar graficamente o desempenho do classificador, relacionando sensibilidade e especificidade. A eficiência do classificador é observada na medida em que a curva se aproxima ao classificador ideal: taxa de falso positivo igual a zero e a taxa de verdadeiro positivo igual a 1, correspondente ao ponto superior esquerdo do gráfico, coordenadas (0,1) e a área sob a curva ROC (AUC, do inglês *Area Under Curve*) idealmente próximo de 1. Assim, o classificador é efetivo se $AUC > 0,5$ e a curva ROC estiver acima da linha tracejada diagonal, a qual representa um classificador aleatório.

4. RESULTADOS E DISCUSSÕES

Os testes iniciais foram realizados utilizando as 16 variáveis independentes listadas no Quadro 2. No entanto, para melhoria do desempenho computacional foram excluídas as variáveis independentes de menor relevância para o modelo: tipo da entidade, nível da entidade, tipo de fornecedor, mês que o contrato foi assinado e ano que o contrato foi assinado. O modelo final utilizou 11 variáveis para previsão das multas, indicando fraude.

Os resultados do modelo RF para o conjunto de variáveis selecionadas, considerando o ano de 2022, constam nas Tabelas 1 a 4. Os dados foram analisados mensalmente e anualmente através da matriz de confusão (Tabelas 1 e 2) e das medidas de acurácia, precisão, *recall* e *F1 Score* (Tabelas 3 e 4).

As matrizes de confusão obtidas por meio do modelo encontram-se nas Tabela 1 e Tabela 2. Observa-se que para a análise anual, o erro do tipo I (quando o modelo identifica um falso positivo) e o erro do tipo II (quando o modelo deixa de identificar um negativo) não afetam significativamente o modelo, considerando que a acurácia foi de 0,78 e o *F1 Score* foi maior que 0,74. No que se relaciona às matrizes mensais, a média geral da acurácia e do *F1 Score* foram superiores a 0,7. Ao pormenorizar a análise para cada mês, temos situações como o caso do mês de julho. Neste caso, as métricas indicam que o modelo tem uma boa sensibilidade (*recall*), identificando corretamente a maioria dos casos positivos, mas a precisão é menor (0,66), sugerindo que há uma quantidade significativa de falsos positivos. A acurácia geral é razoável, mas pode haver espaço para melhorias, especialmente em relação à precisão.

TABELA 1: Matriz de confusão mensal (amostra de 10%)

Janeiro	Predito positivo	Predito negativo	Fevereiro	Predito positivo	Predito negativo	Março	Predito positivo	Predito negativo
Real positivo	590	54	Real positivo	2927	2702	Real positivo	4603	2849
Real negativo	0	595	Real negativo	548	4964	Real negativo	698	6752
Abril	Predito positivo	Predito negativo	Mai	Predito positivo	Predito negativo	Junho	Predito positivo	Predito negativo
Real positivo	2202	458	Real positivo	17155	6785	Real positivo	4854	1124
Real negativo	17	2658	Real negativo	4462	19609	Real negativo	76	5927
Julho	Predito positivo	Predito negativo	Agosto	Predito positivo	Predito negativo	Setembro	Predito positivo	Predito negativo
Real positivo	1694	1665	Real positivo	535	149	Real positivo	1851	1404
Real negativo	119	3221	Real negativo	17	618	Real negativo	665	2565
Outubro	Predito positivo	Predito negativo	Novembro	Predito positivo	Predito negativo	Dezembro	Predito positivo	Predito negativo
Real positivo	14357	7035	Real positivo	28604	9667	Real positivo	4457	849
Real negativo	4391	16933	Real negativo	20546	17922	Real negativo	523	4742

Fonte: Elaboração própria.

TABELA 2: Matriz de confusão anual (amostra de 10%)

2022	Predito positivo	Predito negativo
Real positivo	114501	60487
Real negativo	18011	156264

Fonte: Elaboração própria.

A Tabela 4 mostra os resultados do modelo para uma amostra de 10% dos dados, considerando as informações de todo o período, ou seja, o ano de 2022. A acurácia média foi de 0,78, ligeiramente superior à média mensal (ver Tabela 3). Já o *FI Score* do modelo apresentou valores de 0,74 para irregularidade não detectada e de 0,8 para irregularidade detectada.

As informações consignadas na Tabela 3, foram geradas com amostra de 10% dos dados. A acurácia média foi de 0,76, sendo que o maior valor foi de 0,96 em Janeiro e o menor valor foi de 0,61 para o mês de Novembro. Dessa forma, pode-se afirmar que o modelo acertou as previsões realizadas em 76% das vezes.

TABELA 3: Resultados do modelo RF por mês (amostra de 10%)

Mês	Situação	Precision	Recall	F1-Score	Support
Janeiro	Irregularidade não detectada (0)	1	0,92	0,96	644
	Irregularidade detectada (1)	0,92	1	0,96	595
	Acurácia			0,96	1239
Fevereiro	Irregularidade não detectada (0)	0,84	0,52	0,64	5629
	Irregularidade detectada (1)	0,65	0,9	0,75	5512
	Acurácia			0,71	11141
Março	Irregularidade não detectada (0)	0,87	0,62	0,72	7452
	Irregularidade detectada (1)	0,7	0,91	0,79	7450
	Acurácia			0,76	14902
Abril	Irregularidade não detectada (0)	0,99	0,83	0,9	2660
	Irregularidade detectada (1)	0,85	0,99	0,92	2675
	Acurácia			0,91	5335
Maió	Irregularidade não detectada (0)	0,79	0,72	0,75	23940
	Irregularidade detectada (1)	0,74	0,81	0,78	24071
	Acurácia			0,77	48011
Junho	Irregularidade não detectada (0)	0,98	0,81	0,89	5978
	Irregularidade detectada (1)	0,84	0,99	0,91	6003
	Acurácia			0,9	11981
Julho	Irregularidade não detectada (0)	0,93	0,5	0,66	3359
	Irregularidade detectada (1)	0,66	0,96	0,78	3340
	Acurácia			0,73	6699
Agosto	Irregularidade não detectada (0)	0,97	0,78	0,87	684
	Irregularidade detectada (1)	0,81	0,97	0,88	635
	Acurácia			0,87	1319
Setembro	Irregularidade não detectada (0)	0,74	0,57	0,64	3255
	Irregularidade detectada (1)	0,65	0,79	0,71	3230
	Acurácia			0,68	6485
Outubro	Irregularidade não detectada (0)	0,77	0,67	0,72	21392
	Irregularidade detectada (1)	0,71	0,79	0,75	21324
	Acurácia			0,73	42716
Novembro	Irregularidade não detectada (0)	0,58	0,75	0,65	38271
	Irregularidade detectada (1)	0,65	0,47	0,54	38468
	Acurácia			0,61	76739
Dezembro	Irregularidade não detectada (0)	0,89	0,84	0,87	5306
	Irregularidade detectada (1)	0,85	0,9	0,87	5265
	Acurácia			0,87	10571
Média	Irregularidade não detectada (0)	0,88	0,735	0,735	94630
	Irregularidade detectada (1)	0,725	0,905	0,785	94497
	Acurácia			0,765	189127

Fonte: Elaboração própria

TABELA 4: Resultados do modelo RF (amostra de 10%)

Ano	Situação	Precision	Recall	F1-Score	Support
2022	Irregularidade não detectada (0)	0,86	0,65	0,74	174988
	Irregularidade detectada (1)	0,72	0,9	0,8	174275
	Acurácia			0,78	349263

Fonte: Elaboração própria

Os valores de acurácia foram semelhantes aos obtidos por Henrique, Sobreiro e Kimura (2020) igual 75,58% e inferiores aos obtidos por Sá, Pessanha e Alves (2024) igual a 95,83%.

Além disso, em relação ao *F1 Score*, o modelo apresentou valores médios de 0,72 para irregularidade não detectada e de 0,79 para irregularidade detectada. Esses valores são indicativos de que o modelo está fazendo boas previsões tanto para a classe positiva quanto para a classe negativa. Para classe negativa, o modelo tem dificuldades significativas em corretamente identificar casos que não são irregulares, como indicado pelo baixo *recall*. Isso sugere que muitos casos legítimos podem ser erroneamente sinalizados como irregulares, potencialmente causando interrupções e investigações desnecessárias. Para classe positiva, o modelo foi eficaz em detectar irregularidades, com um alto *recall*, o que é positivo para a detecção de fraude. No entanto, a precisão mais baixa implica um número considerável de falsos positivos, que também podem levar a custos operacionais adicionais e possíveis inconvenientes.

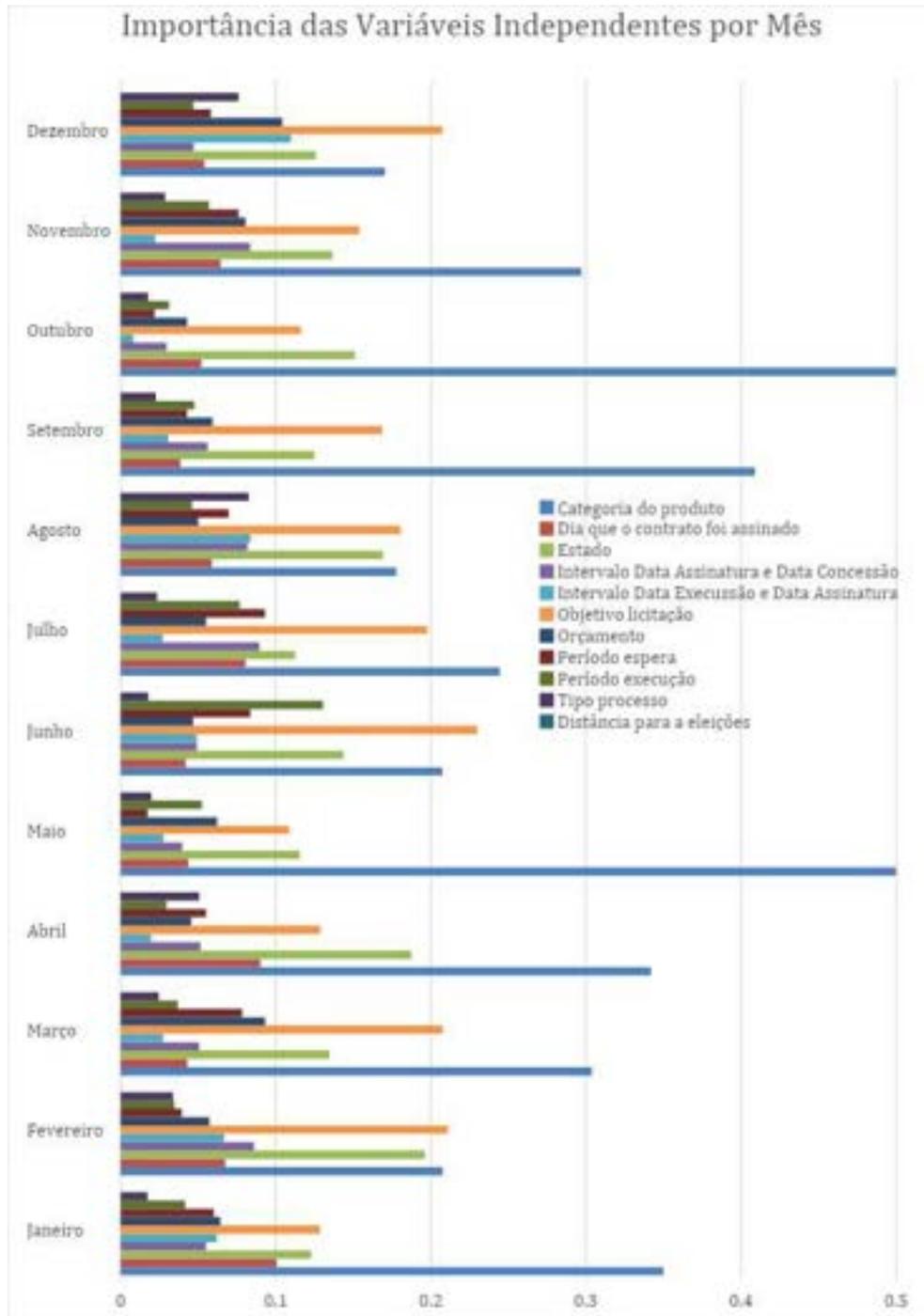
O menor valor obtido para *F1 Score* foi de 0,54 para irregularidade detectada no mês de Novembro. Registra-se que este foi o único valor menor que 0,6 para este coeficiente. Em novembro, nota-se uma queda significativa na precisão para casos não irregulares e um aumento no *recall* para a mesma classe, indicando que o modelo passou a identificar mais eficazmente os casos legítimos como tais, mas com um aumento correspondente em falsos positivos. Para casos de irregularidades, houve uma redução considerável no *recall* (de 1,00 para 0,47), o que significa que o modelo deixou de identificar muitos casos reais de irregularidades. De fato, nota-se uma mudança no padrão do modelo de setembro a novembro em relação aos meses precedentes.

O último trimestre do ano é tradicionalmente um período de intensa atividade em licitações públicas no Brasil. Muitas agências buscam utilizar os orçamentos restantes para garantir que os fundos não sejam desperdiçados ou reduzidos no próximo ciclo fiscal. No ano de 2022 isso pode ainda ter sido intensificado pelas eleições presidenciais, marcadas pela forte polarização política. O resultado das eleições no final de outubro e a mudança de governo pode ter influenciado uma mudança no padrão dos dados de licitação, refletindo-se em uma piora do

modelo de previsão observada especialmente no mês de novembro. Outra possível explicação que pode ter influenciado na baixa performance do modelo neste mês, foi o volume de dados utilizado consideravelmente superior correspondente a 32% dos dados totais, o que pode ter causado um subajuste do modelo.

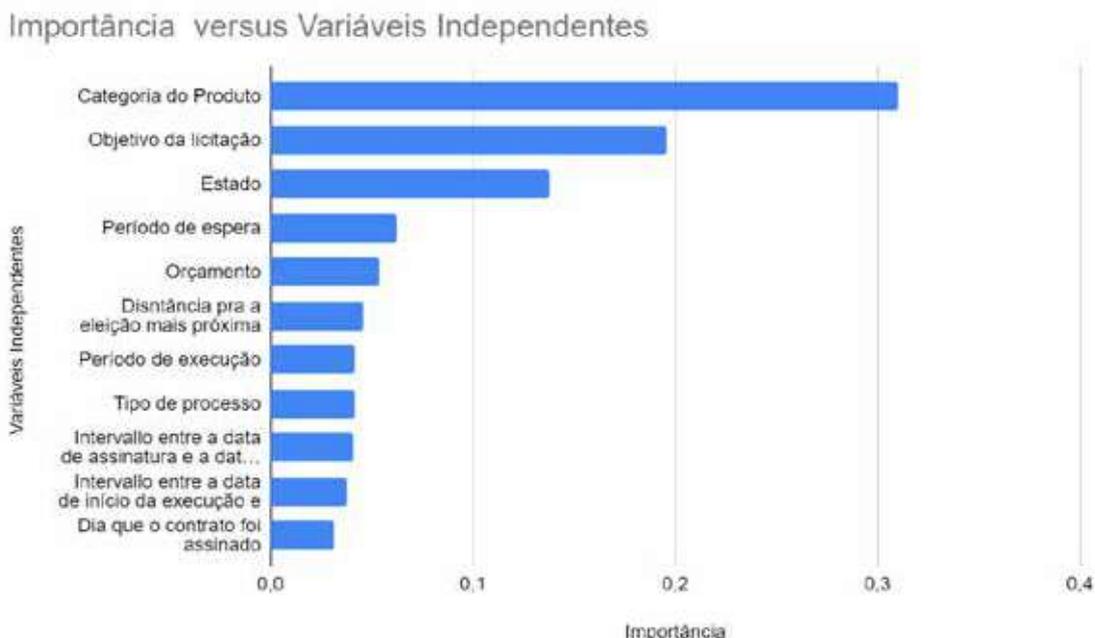
Outro resultado obtido do modelo RF é uma lista com a importância das variáveis independentes (Figura 1 e Figura 2). Para a análise anual, Categoria do Produto, Objetivo da Licitação e Estado são as variáveis mais contribuem para a precisão do modelo, sendo que somadas, respondem por 64,40%. Excluindo-se as três variáveis com maior contribuição para o modelo, observa-se um certo balanceamento entre a importância das demais variáveis, sendo que apresentou o melhor resultado obteve 0,06 e a que obteve o pior resultado foi 0,03. Deste segundo grupo as variáveis mais importantes são Orçamento (0,06), o Período de Espera (0,05) e a Distância para a eleição mais próxima (0,05).

FIGURA 1: Importância das variáveis independentes por mês (amostra de 10%)



Fonte: Elaboração própria.

Figura 2: Importância das variáveis independentes anual (amostra de 10%)



Fonte: Elaboração própria.

Embora este estudo tenha utilizado variáveis do trabalho de Gallego, Rivero e Martínez (2021), eles utilizaram modelos de classificação diferentes, comparando o lasso e GBM, com dados da Colômbia. Eles obtiveram como variáveis independentes chaves as características típicas dos contratos como seu tamanho (medido pelo orçamento estimado e duração), atrasos na implementação, tempo antes da próxima eleição, e padrões geográficos e específicos do setor. Embora no geral a importância das variáveis tenha sido similar entre os dois estudos, em Gallego, Rivero e Martínez (2021) o Orçamento foi a variável mais relevante, enquanto esse estudo obteve a Categoria de Produto como mais significativa. Isso pode indicar que contextos geográficos diferentes com fatores econômicos, políticos e culturais específicos podem afetar a relevância das variáveis. Ou a diferença pode estar simplesmente relacionada as diferenças nos modelos utilizados, sendo algo que deve ser aprofundado na escolha do método, sugerindo que determinados modelos de classificação podem ser melhor para detectar certos contextos.

Sá, Pessanha e Alves (2024) utilizando RF com dados do Rio de Janeiro, obteve resultados semelhantes ao de Gallego, Rivero e Martínez (2021), onde a principal variável foi o ‘Valor dos contratos registrados para o fornecedor do período’, seguido por ‘Quantidade de contratos registrados para o fornecedor do período’ e capital social.

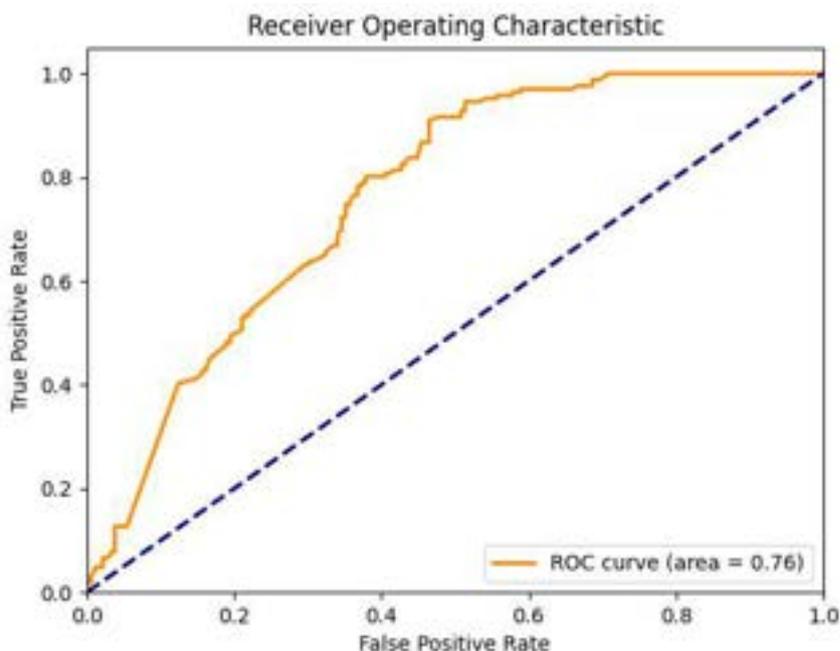
A análise mensal (Figura 1) mostrou haver variações na importância das variáveis independentes ao longo do ano. Por exemplo, a variável Orçamento que veio em quarto em importância na análise anual, com exceção dos meses de Agosto e Dezembro, foi menos importante

em todos os outros meses.

Na análise mensal, a variável 'Distância para a eleição mais próxima' não exibiu relevância em nenhum mês estudado. Isso se deve ao fato de que, dentro de cada mês, a distância temporal até a próxima eleição é constante para todos os dados analisados, eliminando variações que poderiam evidenciar sua importância no modelo.

A área abaixo da curva (AUC - *area under the curve*) alcançada para uma amostra de teste com 0,01% dos dados anuais foi de 76%, ligeiramente superior ao obtido por Henrique, Sobreiro e Kimura (2020) (74,92%). A curva ROC obtida é mostrada na Figura 3.

FIGURA 3: Curva ROC para amostra de 0,01% dos dados anuais



Fonte: Elaboração própria.

5. CONCLUSÃO

A informatização dos processos de licitação trouxe maior transparência e eficiência, reduzindo custos operacionais e acelerando o processo de contratação. Além disso, promoveu uma concorrência mais ampla. No entanto, esse avanço também impõe aos órgãos de auditoria a necessidade de adotar novas ferramentas para a detecção de fraudes, adaptando-se as inovações tecnológicas no setor e ao volume massivo de dados envolvido. Assim, existe uma tendência crescente de aplicação de técnicas de IA e MLs na fiscalização dos sistemas eletrônicos de contratações públicas.

Modelos de classificação permitem prever a ocorrência de fraudes, baseados em padrões pré-identificados em um conjunto de dados rotulado, e apontar as principais variáveis indicadoras de irregularidades, direcionando os esforços de investigação. Dentre essas técnicas, as características do modelo RF são particularmente adequadas para esta aplicação, devido à sua robustez, precisão, capacidade de manejar a complexidade dos dados (grande volume, alta dimensionalidade e valores ausentes), além de sua simplicidade de implementação. Destaca-se também por fornecer uma medida de importância das variáveis. Vários estudos comparativos entre os principais algoritmos comprovaram sua eficiência no contexto de detecção de fraudes em licitações brasileiras (HENRIQUE; SOBREIRO; KIMURA, 2020; SÁ; PESSANHA; ALVES, 2024) e outros ainda indicaram que era o melhor modelo de classificação (FARIA, 2023; SILVA, 2022).

O modelo RF foi implementado utilizando dados do Portal Transparência da Controladoria-Geral da União, do período de 01/01/2022 a 31/12/2022. Foi utilizado um conjunto de 11 variáveis independentes: orçamento, objetivo da licitação, tipo de processo, dia que o contrato foi assinado, período de execução, período de espera, intervalo entre a data de assinatura e a data de concessão, intervalo entre a data de início da execução e a data de assinatura, distância para a eleição mais próxima, categoria do produto, estado. A variável dependente foi definida como multas recebidas pelas empresas. Os dados foram divididos em 70% para treinamento e 30% para teste. O classificador foi implementado com 100 árvores.

O classificador implementado teve um desempenho eficiente com *F1 Score* mensal médio de 78,5% e anual de 80%. Destacou-se pela capacidade de detectar a maioria dos casos reais de fraude, com *recall* mensal médio e anual de 90%. As variáveis de maior importância para o modelo foram Categoria do Produto, Objetivo da Licitação e Estado para ambas as análises, mensal e anual. No entanto, a importância dessas variáveis flutuou significativamente ao longo dos meses.

Observou-se uma mudança de padrão significativa entre os meses, que deve ser investigada mais profundamente. Para casos de irregularidades, houve uma redução considerável nos parâmetros de qualidade com valores máximo em Janeiro e mínimo em Novembro (variação no *recall* de 1,00 para 0,47, respectivamente, e variação no *F1 Score* de 0,96 para 0,54, respectivamente). Apesar que a variável ‘distância para a eleição mais próxima’ ser somente a sexta em nível de importância, acredita-se que as eleições presidenciais de 2022 tenham influenciado significativamente na mudança de padrão.

A pesquisa foi limitada por restrições computacionais. O grande volume de dados aumen-

tou significativamente o tempo de processamento, exigindo a divisão dos dados por mês para realizar a análise. O processamento anual demandou um computador com pelo menos 32GB de memória RAM e, mesmo assim, levou aproximadamente 1 hora para concluir o processamento dos dados. Ademais, somente a variável dependente de multas recebidas pelas empresas não captura todos os casos de fraudes existentes. Investigações em andamento ou denúncias são possíveis variáveis dependentes a serem incorporadas no modelo em trabalhos futuros para ampliar o espectro de detecção de fraudes. A importância de outras variáveis independentes também deverá ser avaliada. Por fim, análises futuras deverão incluir outros anos para comparação de desempenho e refinamento do modelo.

O modelo RF desenvolvido tem potencial de aplicação prática além da academia, oferecendo aos gestores públicos uma ferramenta eficaz na detecção de possíveis desvios de conduta no processo de licitação.

REFERÊNCIAS BIBLIOGRÁFICAS

ABREU, B. M.; PEREIRA, T. H. S.; GOMES-JR, L. Detecção de Fraudes em Licitações Públicas: Uma Comparação de Modelos de Detecção de Anomalias. In: ESCOLA REGIONAL DE BANCO DE DADOS (ERBD), 2024, Farroupilha/RS. Anais [...]. Porto Alegre: **Sociedade Brasileira de Computação**, p. 81-90, 2024. DOI: <https://doi.org/10.5753/erbd.2024.238821>.

ADOBOR, H.; YAWSON, R. The promise of artificial intelligence in combating public corruption in the emerging economies: A conceptual framework. **Science and Public Policy**, v. 50, n. 3, p. 355-370, 2023. DOI: <https://doi.org/10.1093/scipol/scac068>.

ALDANA, A.; FALCÓN-CORTÉS, A.; LARRALDE, H. A machine learning model to identify corruption in México's public procurement contracts. arXiv preprint arXiv:2211.01478, 2022.

AMORIM, V. A. J. Licitações e contratos administrativos: teoria e jurisprudência. Brasília: Senado Federal, 2020. Disponível em: <https://bibliotecadigital.stf.jus.br/xmlui/bitstream/handle/123456789/2461/1182434.pdf?sequence=1&isAllowed=y>. Acesso em: 24 jul. 2024.

BIAGINI, E. Ferramenta de IA já evitou gastos de R\$ 11 bi em licitações suspeitas, diz ministro da CGU. **Agência Gov**, 2024. Disponível em: <https://agenciagov.ebc.com.br/noticias/ferramenta-de-ia-da-cgu-ja-economizou-r-11-bilhoes-em-gastos-publicos-em-licitacoes-suspeitas-diz-ministro>>. Acesso em: 02 ago. 2024.

BRASIL. Constituição (1988). Constituição da República Federativa do Brasil de 1988. Promulgada em 05 de outubro de 1988. Brasília, DF: Senado Federal, Centro Gráfico, 1988.

BRASIL. Lei nº 8.666, de 21 de junho de 1993. Estabelece normas para licitações e contratos da Administração Pública e dá outras providências. Diário Oficial da União, Brasília, DF, 22 jun. 1993. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/18666cons.htm.

BRASIL. Lei nº 10.520/2002, de 17 de julho de 2002. Institui, no âmbito da União, Estados, Distrito Federal e Municípios, nos termos do art. 37, inciso XXI, da Constituição Federal, modalidade de licitação denominada pregão, para aquisição de bens e serviços comuns, e dá outras

providências. Diário Oficial da União, Brasília, DF, 17 jul. 2002. Disponível em: https://www.planalto.gov.br/ccivil_03/LEIS/2002/L10520.htm?origin=instituicao.

BRASIL. Lei nº 12.462, de 4 de agosto de 2011. Institui o Regime Diferenciado de Contratações Públicas – RDC; altera a Lei nº 10.683, de 28 de maio de 2003, que organiza a Presidência da República e os Ministérios, e dá outras providências. Diário Oficial da União, Brasília, DF, 5 ago. 2011. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/112462.htm.

BRASIL. Lei nº 12.846, de 1º de agosto de 2013. Dispõe sobre a responsabilização administrativa e civil de pessoas jurídicas pela prática de atos contra a administração pública, nacional ou estrangeira, e dá outras providências. Diário Oficial da União, Brasília, DF, 2 ago. 2013. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2013/lei/112846.htm.

BRASIL. Lei nº 14.133, de 1º de abril de 2021. Lei de Licitações e Contratos Administrativos. Diário Oficial da União, Brasília, DF, 1 abr. 2021. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2021/lei/114133.htm.

BREIMAN, Leo. Random forests. **Machine learning**, v. 45, p. 5-32, 2001. <https://link.springer.com/article/10.1023/A:1010933404324>

CONTROLADORIA GERAL DA UNIÃO. Alice - Analisador de Licitações, Contratos e Editais. Disponível em: <https://www.gov.br/cgu/pt-br/assuntos/auditoria-e-fiscalizacao/alice>. Acesso em: 02 ago. 2024.

DECAROLIS, F.; GIORGIANTONIO,. Corruption red flags in public procurement: new evidence from Italian calls for tenders. **EPJ Data Science**, v. 11, n. 1, p. 16, 2022. <https://doi.org/10.1140/epjds/s13688-022-00325-x>

FARIA, G. A. Fraudes em compras governamentais: Detecção com Aprendizado de Máquina. 2023. Trabalho de Conclusão de Curso (Graduação em Administração Pública) - Escola de Administração Pública, Centro de Ciências Jurídicas e Políticas, Universidade Federal do Estado do Rio de Janeiro, Rio de Janeiro, 2023.

GADOUR, M. M. E. Corruption in public procurement: Can e-procurement and artificial intelligence make a difference in Africa? **QScience Connect**, v. 2024, n. 1-Thesis, p. 2, 2024. DOI: <https://doi.org/10.5339/connect.2024.spt.2>.

GALLEGO, J.; RIVERO, G.; MARTÍNEZ, J. Preventing rather than punishing: An early warning model of malfeasance in public procurement. **International Journal of Forecasting**, v. 37, n. 1, p. 360-377, 2021. <http://dx.doi.org/10.1016/j.ijforecast.2020.06.006>.

GUNASEGARAN, M.; BASIRUDDIN, R.; RIZAL, A. M. Detecting and Preventing Fraud in e-procurement of Public Sector: A Review, Synthesis and Opportunities for Future Research. **International Journal of Academic Research in Business and Social Sciences**, v. 13, n.1, p. 1444 – 1463, 2023.

HENRIQUE, B. M.; SOBREIRO, V. A.; KIMURA, H. Contracting in Brazilian public administration: A machine learning approach. **Expert Systems**. 2020. <https://doi.org/10.1111/exsy.12550>

LIMA, M. C.; SILVA, R. C.; MENDES, F. L. S.; CARVALHO, L. R.; ARAUJO, A.; VIDAL, F. B. Inferring about fraudulent collusion risk on Brazilian public works contracts in official texts using a Bi-LSTM approach. In: COHN, T.; HE, Y.; LIU, Y. (Ed.). Findings of the Association for Computational Linguistics: EMNLP 2020. Online: **Association for Computational Linguistics**, p. 1580-1588, 2020. Disponível em: <https://aclanthology.org/2020.findings-emnlp.143>. Acesso em: 22 jul. 2024. DOI: 10.18653/v1/2020.findings-emnlp.143.

LOPES, M. A. Aplicação de aprendizado de máquina na detecção de fraudes públicas. 2019. Dissertação (Mestrado em Administração) - Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, 2019. doi:10.11606/D.12.2020.tde-10022020-174317. Acesso em: 2024-07-22.

MAVIDIS A, FOLINAS D. From Public E-Procurement 3.0 to E-Procurement 4.0; A Critical Literature Review. **Sustainability**, v. 14, n. 18, 2022. <https://doi.org/10.3390/su141811252>

NAI, Roberto et al. Public procurement fraud detection and artificial intelligence techniques: a literature review. In: **Companion Proceedings of the 23rd International Conference on Knowledge Engineering and Knowledge Management**. CEUR-WS, p. 1-13, 2022. https://iris.unito.it/bitstream/2318/1888876/1/KM4LAW22_preprint.pdf

OLIVEIRA, G. P. et al. Ranqueamento de Licitações Públicas a partir de Alertas de Fraude. In: **Anais do XII Brazilian Workshop on Social Network Analysis and Mining**. SBC, 2023. p. 1-12. <https://doi.org/10.5753/brasnam.2023.232105>

REJEB, A.; REJEB, K.; APPOLLONI, A.; KAYIKCI, Y.; IRANMANESH, M. The landscape of public procurement research: A bibliometric analysis and topic modelling based on Scopus. **Journal of Public Procurement**, v. 23, n. 2, 2023. <https://doi.org/10.1108/JOPP-06-2022-0031>

RIBEIRO, C. G.; INÁCIO, E. O mercado de compras governamentais brasileiro (2006-2017): Mensuração e análise. Brasília: Instituto de Pesquisa Econômica Aplicada (IPEA), 2019. Texto para Discussão, No. 2476. Disponível em: <<https://hdl.handle.net/10419/211431>>. Acesso em: 02 ago. 2024.

SÁ, T. A.; PESSANHA, J. F. M.; ALVES, F. J. S. Métodos de classificação supervisionada aplicados à identificação de fraudes de fornecedores. **CONTABILOMETRIA - Brazilian Journal of Quantitative Methods Applied to Accounting, Monte Carmelo**, v. 11, n. 2, p.125-146, 2024.

SCIKIT-LEARN. **RandomForestClassifier**. Disponível em: <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html#sklearn.ensemble.RandomForestClassifier>. Acesso em: 02 ago. 2024.

SECOM TCU. **Chatbot do TCU fornece certidões pelo whatsapp**. 2020. Disponível em: <https://portal.tcu.gov.br/imprensa/noticias/chatbot-do-tcu-fornece-certidoes-pelo-whatsapp.htm>. Acesso em: 02 ago. 2024.

SICILIANI, L.; TACCARDI, V.; BASILE, P.; DI CIANO, M.; LOPS, P. AI-based decision support system for public procurement. **Information Systems**, v. 119, 2023.

SILVA, G. Prevenção de irregularidades em contratos públicos: uma análise utilizando técnicas de Machine Learning. 2022. 91 f., il. Trabalho de Conclusão de Curso (Bacharelado em Administração) — Universidade de Brasília, Brasília, 2022.

SUPREMO TRIBUNAL FEDERAL. **Projeto Victor avança em pesquisa e desenvolvimento para identificação dos temas de repercussão geral**. 2021. Disponível em: <<https://portal.stf.jus.br/noticias/verNoticiaDetalhe.asp?idConteudo=471331&ori=1>>. Acesso em: 02 ago. 2024.

SUPREMO TRIBUNAL FEDERAL. **STF finaliza testes de nova ferramenta de Inteligência Artificial**. 2023a. Disponível em: <<https://portal.stf.jus.br/noticias/verNoticiaDetalhe.asp?idConteudo=507120&ori=1>>. Acesso em: 02 ago. 2024.

TRIBUNAL DE CONTAS DA UNIÃO. **Guia de uso de inteligência artificial generativa no Tribunal de Contas da União (TCU)**. Disponível em: <<https://portal.tcu.gov.br/data/files/42/F7/91/4B/B59019105E366F09E18818A8/Guia%20de%20uso%20de%20IA%20generativa%20no%20TCU.pdf>>. Acesso em: 02 ago. 2024.

VELASCO, R. B.; CARPANESE, I.; INTERIAN, R.; PAULO NETO, O.C.G.; RIBEIRO, C.C. A decision support system for fraud detection in public procurement. **International Transactions Inoperational Research**, v. 28, p. 27-47, 2021. <https://doi.org/10.1111/itor.12811>

ZAGO, M. F. **Poder de compra estatal como instrumento de políticas públicas?** Brasília: Escola Nacional de Administração Pública, 2018.