



Universidade Federal de Campina Grande
Centro de Engenharia Elétrica e Informática
Coordenação de Pós-Graduação em Ciência da Computação

Talita Lôbo de Menezes

**Estratégias de Modelagem e Características Influentes na
Composição do Risco de Rescisão em Contratos Públicos**

Campina Grande - PB
2025

Universidade Federal de Campina Grande
Centro de Engenharia Elétrica e Informática
Coordenação de Pós-Graduação em Ciência da Computação

Estratégias de Modelagem e Características
Influentes na Composição do Risco de Rescisão em
Contratos Públicos

Talita Lôbo de Menezes

Tese submetida à Coordenação do Curso de Pós-Graduação em Ciência da Computação da Universidade Federal de Campina Grande - Campus I como parte dos requisitos necessários para obtenção do grau de Doutor em Ciência da Computação.

Área de Concentração: Ciência da Computação
Linha de Pesquisa: Sistemas de Computação

Dr. Nazareno Ferreira de Andrade (Orientador)
Dr. Fábio Jorge Almeida de Moraes (Co-orientador)

Campina Grande, Paraíba, Brasil
©Talita Lôbo de Menezes, 15/02/2025

M543e Menezes, Talita Lôbo.
Estratégias de modelagem e características influentes na composição do risco de rescisão em contratos públicos / Talita Lôbo Menezes – Campina Grande, 2025.
154 f. il. color.

Tese (Doutorado em Ciência da Computação) - Universidade Federal de Campina Grande, Centro de Engenharia Elétrica e Informática, 2025.
"Orientação: Prof. Dr. Nazareno Ferreira de Andrade. Dr. Fábio Jorge Almeida de Moraes."

Referências.
1. Sistemas de Computação. 2. Aprendizagem de Máquina. 3. Estimação de Risco. 4. Engenharia de Características. I. Andrade, Nazareno Ferreira de. II. Moraes, Fábio Jorge Almeida de. III. Título.

CDU 004.67(043)

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECARIA ITAPUANA SOARES DIAS GONÇALVES CRB-15/093



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE CAMPINA GRANDE
POS-GRADUACAO EM CIENCIA DA COMPUTACAO
Rua Aprígio Veloso, 882, Edifício Telmo Silva de Araújo, Bloco CG1, - Bairro Universitário, Campina Grande/PB, CEP 58429-900
Telefone: 2101-1122 - (83) 2101-1123 - (83) 2101-1124
Site: <http://computacao.ufcg.edu.br> - E-mail: secp@computacao.ufcg.edu.br

FOLHA DE ASSINATURA PARA TESES E DISSERTAÇÕES

TALITA LÔBO DE MENEZES

ESTRATÉGIAS DE MODELAGEM E CARACTERÍSTICAS INFLUENTES NA COMPOSIÇÃO DO RISCO DE RESCISÃO EM CONTRATOS PÚBLICOS

Tese apresentada ao Programa de Pós-Graduação em Ciência da Computação como pré-requisito para obtenção do título de Doutor em Ciência da Computação.

Aprovada em: 14/03/2025

Prof. Dr. **Nazareno Ferreira de Andrade**, UFCG, Orientador

Prof. Dr. **Fábio Jorge Almeida Moraes**, UFCG, Orientador

Prof. Dr. **Leandro Balby Marinho**, UFCG, Examinador Interno

Prof. Dr. **José Antônio Beltrão Moura**, UFCG, Examinador Interno

Prof. Dr. **Diego Furtado Silva**, USP, Examinador Externo

Prof. Dr. **Flavio Vinicius Diniz de Figueiredo**, UFMG Examinador Externo



Documento assinado eletronicamente por **FABIO JORGE ALMEIDA MORAIS, PROFESSOR(A) DO MAGISTERIO SUPERIOR**, em 17/03/2025, às 10:23, conforme horário oficial de Brasília, com fundamento no art. 8º, caput, da [Portaria SEI nº 002, de 25 de outubro de 2018](#).



Documento assinado eletronicamente por **Flavio Vinicius Diniz de Figueiredo, Usuário Externo**, em 17/03/2025, às 14:36, conforme horário oficial de Brasília, com fundamento no art. 8º, caput, da [Portaria SEI nº 002, de 25 de outubro de 2018](#).



Documento assinado eletronicamente por **JOSE ANTAO BELTRAO MOURA, PROFESSOR 3 GRAU**, em 19/03/2025, às 14:08, conforme horário oficial de Brasília, com fundamento no art. 8º, caput, da [Portaria SEI nº 002, de 25 de outubro de 2018](#).



Documento assinado eletronicamente por **LEANDRO BALBY MARINHO, PROFESSOR(A) DO MAGISTERIO SUPERIOR**, em 20/03/2025, às 10:46, conforme horário oficial de Brasília, com fundamento no art. 8º, caput, da [Portaria SEI nº 002, de 25 de outubro de 2018](#).



Documento assinado eletronicamente por **NAZARENO FERREIRA DE ANDRADE, Usuário Externo**, em 03/04/2025, às 17:53, conforme horário oficial de Brasília, com fundamento no art. 8º, caput, da [Portaria SEI nº 002, de 25 de outubro de 2018](#).



Documento assinado eletronicamente por **Diego Furtado Silva, Usuário Externo**, em 08/04/2025, às 09:24, conforme horário oficial de Brasília, com fundamento no art. 8º, caput, da [Portaria SEI nº 002, de 25 de outubro de 2018](#).



A autenticidade deste documento pode ser conferida no site <https://sei.ufcg.edu.br/autenticidade>, informando o código verificador **5303484** e o código CRC **BD46257C**.

Resumo

Contratos públicos são instrumentos utilizados para a formalização de vínculos entre a administração pública e empresas contratadas. Com o uso desses instrumentos pelas administrações públicas federal, estaduais e municipais para a aquisição de bens e serviços, com o objetivo de suprir as demandas da população, o volume de contratos firmados representa um desafio para os órgãos públicos e a sociedade civil no pleno exercício do controle e acompanhamento da execução desses contratos. Nesse contexto, o uso da aprendizagem de máquina na seleção de casos prioritários para fiscalização tem se mostrado uma estratégia promissora. Este trabalho de tese avaliou a eficácia prática de uma ferramenta atualmente em uso para a priorização de empresas com comportamentos de risco. Nessa avaliação, os resultados mostram que a priorização realizada pela ferramenta é ineficaz, especialmente ao considerar o número de empresas falsamente classificadas como de alto risco, reforçando o enquadramento do problema da estimação de risco em nível do contrato. Em seguida, com um conjunto inédito de rótulos em nível de contrato, foi proposto e avaliado um modelo de previsão de risco de contratos municipais treinado a partir de características do próprio contrato, da empresa contratada e de seu histórico de fornecimento, além da avaliação de estratégias de mitigação do desbalanceamento entre classes de risco. O resultado da modelagem evidencia a viabilidade da estimação de risco em nível de contrato, com performance superior à de um classificador aleatório para o conjunto de dados considerado. Entretanto, não há evidências de que o uso de técnicas de balanceamento artificial de classes nesse contexto melhore a capacidade preditiva do modelo em relação ao balanceamento por pesos durante o aprendizado. A partir desse modelo base apresentado, outras características preditoras foram experimentadas: características latentes sobre o objeto do contrato e a característica econômica do PIB *per capita* dos municípios contratantes. Como resultado, observa-se que, ao utilizar características latentes, a performance do modelo é equivalente ou superior à do modelo base. Já com o uso do PIB *per capita*, é provável que o classificador obtenha vantagem em relação ao modelo base. Com base no modelo de maior AUC-PR, foi fornecida uma visão geral da influência das características preditoras utilizadas no risco de rescisão contratual e da maneira como cada característica contribui para a classificação correta dos contratos. Por fim, ao avaliar o alinhamento do aprendizado e da previsão do

modelo de risco com os pareceres técnicos de justificativa de rescisão contratual feitos por especialistas, observou-se que as características determinantes para o risco na modelagem possuem razoável verossimilhança com as causas reais e práticas das rescisões contratuais.

Palavras-chave: estimacão de risco; contratos públicos; aprendizagem de máquina; engenharia de características.

Abstract

Public contracts are essential instruments for formalizing agreements between public administrations and contracted companies. Federal, state, and municipal governments rely on these contracts to procure goods and services that address public needs. However, the growing volume of contracts presents significant challenges for public agencies and civil society in effectively monitoring and controlling their execution. In this context, machine learning has emerged as a promising strategy for prioritizing high-risk cases for inspection. This dissertation evaluates the practical effectiveness of a tool currently in use for identifying companies with risk-prone behaviors. Initial findings reveal that the tool’s prioritization is ineffective, particularly due to the high number of companies incorrectly classified as high-risk. These results highlight the need to reframe the risk estimation problem at a more actionable level: the contract level. Using a novel dataset of contract-level labels, this study proposes and evaluates a predictive model for assessing municipal contract risk. The model incorporates features derived from the contract itself, the contracted company, and its supply history, while also exploring strategies to address class imbalance in risk categories. The results demonstrate the feasibility of contract-level risk estimation, with the model outperforming a random classifier on the dataset. However, the use of artificial class balancing techniques does not significantly enhance the model’s predictive performance compared to weight balancing during training. Building on this baseline model, additional predictive features were introduced, including semantically latent features related to the contract’s subject matter and the GDP *per capita* of the contracting municipalities. The inclusion of latent features yielded model performance equivalent to or better than the baseline, while the use of GDP *per capita* provided a potential advantage. Based on the model with the highest AUC-PR, the study offers insights into the influence of predictive features on contract termination risk and their contribution to accurate contract classification. Finally, the study assesses the alignment between the model’s risk predictions and expert technical opinions on contract termination justifications. The findings indicate that the key risk factors identified by the model exhibit reasonable consistency with the real-world causes of contract terminations. This alignment underscores the model’s practical

relevance and its potential to support decision-making in public contract management.

Keywords: risk estimation; public contracts; machine learning; feature engineering.

Agradecimentos

Tão bom quanto começar o processo do doutorado é poder concluí-lo. E, apesar da jornada solitária que é fazer pesquisa nesse campo, tive a sorte e o privilégio de contar com excelentes companhias durante esse tempo.

Agradeço a Nazareno e Fábio – orientadores, amigos e exemplos – pelo tempo, pelas discussões, pela confiança e pelo incentivo ao longo desses últimos anos, dentro ou fora da universidade. Naza, sem seus conselhos, nossas conversas, sua inspiração e animação, dificilmente teria vindo tão longe no percurso acadêmico. Fábio, obrigada por ter embarcado nessa conosco, por ter contribuído tanto. Obrigada por me acompanharem desde a graduação e por me apresentarem uma academia com menos ego, mais humanidade e preocupação com o mundo lá fora. Vocês ainda vão levar muito mais gente além!

Aos competentes amigos e colegas do LSD, pelas conversas de corredor durante o almoço, pelas confraternizações e discussões técnicas. Em especial, agradeço a Caetano e Caio pela companhia nas inúmeras madrugadas de trabalho, e a Luiz Augusto, pela sempre disponível consultoria acadêmica à distância. À Cleide, pelo cuidado diário com nossas burocracias e com o ambiente de trabalho.

Ao suporte do laboratório – que me perdoará por não citar todos que já passaram pelo grupo –, pelas consultorias sobre infraestrutura e pelo auxílio sempre que necessário, sem os quais os experimentos desta tese teriam demorado muitíssimo mais.

Ao promotor Octavio Neto, pela disponibilização de dados, pela confiança, pelas discussões, pelas oportunidades e também por todo incentivo desde o mestrado. Aos colegas do Núcleo de Inovação da SEPLAG/MPPB, onde pude aprender muito sobre alguns processos da administração pública, pela hospitalidade e pelo trabalho conjunto.

Aos professores Antão, Diego, Flavio e Leandro pelos valiosos comentários, sugestões e críticas durante a banca de defesa.

À Paloma e Lyanna, da Copin, por todo o suporte nas burocracias administrativas e também pelo incentivo durante esse processo.

Aos amigos, academicamente envolvidos ou não, pela presença, pelas discussões sobre a vida ou sobre a própria tese, pelas correções de texto e distrações sempre que necessário.

Em especial, agradeço à Amanda, Haissa e Natália, pelo tempo, pelas orientações literárias ou correções textuais ágeis.

Agradeço também a cada brasileiro que financiou esta pesquisa.

Por fim, à minha família, em especial Jonas, Tereza, Josefa, Thiago e Saulo. É sempre mais fácil ir quando se tem para onde voltar.

Obrigada.

Conteúdo

1	Introdução	1
1.1	Motivação	3
1.2	Objetivos	7
1.3	Contribuições	8
1.4	Estrutura do Documento	8
2	Fundamentação Teórica	10
2.1	Contratação pública	10
2.1.1	Contrato	12
2.1.2	Desfazimento e mutação contratual	12
2.1.3	SAGRES	13
2.2	Aprendizagem de máquina	14
2.3	Desbalanceamento entre classes	16
2.4	Processamento de linguagem natural	17
2.4.1	Modelagem de Tópicos	17
2.5	XGBoost	18
2.6	Explicabilidade em modelos de aprendizagem	19
2.6.1	Valores SHAPley	20
3	Referências e estado da arte	22
3.1	Estimação de Risco	22
3.1.1	Empresas Licitantes	22
3.1.2	Contratos e Compras	23
3.1.3	Lances e Participantes em Leilão	26

3.1.4 Servidores Públicos	27
3.1.5 Unidades Administrativas	27
3.2 Identificação de Contexto	28
3.3 Considerações	28
4 Materiais e Métodos	31
4.1 Fontes de Dados	31
4.2 Operacionalização do Risco de Rescisão	34
4.2.1 Risco em Empresas	35
4.2.2 Risco em Contratos	36
4.3 Avaliação em Cenários Desbalanceados	36
4.4 Normalização min-max	38
4.5 Infraestrutura	39
5 Estimação de risco de empresas: prática e eficácia	41
5.1 Matriz de risco	42
5.2 Filtragem e rotulagem de dados	42
5.3 Avaliação da matriz de risco	43
5.4 Discussão	45
5.5 Limitações	46
6 Estimação do risco de rescisão em contratos	48
6.1 Método	49
6.1.1 Engenharia de características	49
6.1.2 Extração e tratamento	51
6.1.3 Treino, teste e rótulos	52
6.1.4 Algoritmo e parametrização	52
6.2 Risco do contrato	54
6.3 MR como filtro de contratos	55
6.4 Desbalanço entre classes de risco	56
6.5 Discussão	58

7 Experimentação de outros preditores de risco	63
7.1 Objeto do contrato como característica	63
7.1.1 Extração de tópicos	65
7.1.2 Resultados	67
7.2 Característica econômica dos municípios	68
7.2.1 Resultados	69
7.3 Pesos a partir do valor do contrato	71
7.3.1 Resultados	71
7.4 Impacto de diferentes conjuntos de características	72
7.5 Discussão	75
8 Explicando o risco de rescisão	78
8.1 Influência geral de características	79
8.2 Características relevantes para contratos verdadeiros positivos	87
8.3 Discussão	93
8.4 Limitações	97
9 O risco de rescisão na prática	99
9.1 Contrato I - Construção de unidades básicas de saúde	100
9.2 Contrato II - Confeccção de materiais gráficos	101
9.3 Contrato III - Fornecimento de alimentos	103
9.4 Contrato IV - Prestação de serviços terceirizados	104
9.5 Contrato V - Construção de quadra escolar coberta	105
9.6 Contrato VI - Instalação de itens de serralheria	106
9.7 Discussão	108
10 Conclusões	110
10.1 Limitações	113
10.2 Direções futuras	114
A Resumo - Trabalhos Anteriores	123
B Contratos por Tópico	127

C CNPJs Removidos**130****D *Stopwords*****132**

Lista de Símbolos

ALGL - *Antiga Lei Geral de Licitação*

AM - *Aprendizagem de Máquina*

AP - *Administração Pública*

AUC-PR - *Área sob a Curva de Precisão-Revocação*

AUC-ROC - *Área sob a Curva ROC*

CAFIL - *Cadastro de Fornecedores Impedidos de Licitar e Contratar com a Administração Pública Estadual*

CEAF - *Cadastro de Expulsões da Administração Federal*

CEIS - *Cadastro Nacional de Empresas Inidôneas ou Suspensas*

CGU - *Controladoria Geral da União*

CNCIAI - *Cadastro Nacional de Condenados por Ato de Improbidade Administrativa e por Ato que Implique Inelegibilidade*

CNEP - *Cadastro Nacional de Empresas Punidas*

CV - *Validação Cruzada*

HL - *Hipercubo Latino* IA - *Inteligência Artificial*

INSS - *Instituto Nacional de Seguridade Social*

KNN - *K-Nearest Neighbours*

LAI - *Lei de Acesso à Informação*

MCC - *Coeficiente de Correlação de Matthews*

MPPB - *Ministério Público do Estado da Paraíba*

MSE - *Erro Quadrático Médio*

NLLC - *Nova Lei de Licitações e Contratos*

PLN - *Processamento de Linguagem Natural*

RF - *Receita Federal*

ROC - *Receiver Operating Curve*

SAGRES - *Sistema de Acompanhamento da Gestão dos Recursos da Sociedade*

SIASG SUS - *Sistema Único de Saúde*

SVM - *Support Vector Machines*

TCE - *Tribunal de Contas do Estado*

TCE-PB - *Tribunal de Contas do Estado da Paraíba*

TCU - *Tribunal de Contas da União*

XGBoost - *eXtreme Gradient Boosting*

Lista de Figuras

1.1	Mecanismos e práticas de governança pública organizacional - Fonte: Manual de Licitações e Contratos, Tribunal de Contas da União, 2020, p. 52.	4
4.1	Total de mutações contratuais de interesse ao longo dos anos. Dados: Tramita. Fonte: elaboração própria.	33
5.1	Distribuição do score de risco da MR por classe. Fonte: elaboração própria.	43
5.2	Intervalos de Confiança (AUC-PR) - Matriz de Risco	44
6.1	Classificação do conjunto de testes - Matriz de Confusão	55
6.2	Classificação do conjunto de testes (contratos da MR) - Matriz de confusão	56
6.3	Comparação entre estratégias de mitigação do desbalanço - Matriz de confusão	57
6.4	IC para 95% de confiança - Comparação entre desempenho de modelos	58
6.5	Comparativo de curva de precisão-revoação - GBM vs. XGBoost	60
6.6	IC da diferença entre os modelos <i>SAGRES RF</i> e <i>SAGRES RF (contratos MR)</i>	61
7.1	Ilustração da geração de tópicos por contrato e treinamento do modelo XGBoost	67
7.2	IC para 95% de confiança - Comparação entre desempenho de modelos	67
7.3	Classificação do conjunto de testes - Matriz de confusão	68
7.4	IC para 95% de confiança - Comparação entre desempenho de modelos	69
7.5	IC da diferença entre os modelos <i>SAGRES RF PIBpc</i> e <i>SAGRES RF</i>	70
7.6	Classificação do conjunto de testes - Matriz de confusão	70
7.7	Classificação do conjunto de testes - Matriz de confusão	72
7.8	Impacto da indisponibilidade de conjuntos de características - IC 95% de confiança para AUC-PR	73

7.9	IC da diferença entre os modelos <i>SAGRES RF PIB_{pc}</i> e <i>SAGRES RF Tópicos</i>	74
8.1	Sumário de influência na previsão - Modelo <i>SAGRES + RF + PIB per capita</i>	80
8.2	Dependências parciais – PIB <i>per capita</i>	81
8.3	Dependências parciais – Diferença entre o ano de abertura da empresa e o ano do contrato corrente	83
8.4	Dependências parciais – Valor do contrato em reais (R\$)	84
8.5	Valor do contrato em reais (log10) em relação aos valores SHAPley	84
8.6	Dependências parciais – Razão entre o valor do contrato e valor total recebido	86
8.7	Dependências parciais – Relação entre características (<i>razão entre o valor do contrato e valor total recebido vs. número de propostas feitas pela empresa em licitações</i>)	87
8.8	Gráfico de força individual – Exemplo de contrato verdadeiro positivo . . .	88
8.9	Gráficos de força – Contratos verdadeiros positivos (parte I)	89
8.10	Gráficos de força – Contratos verdadeiros positivos (parte II)	90
9.1	Composição do risco de rescisão – Contrato I	100
9.2	Composição do risco de rescisão – Contrato II	102
9.3	Composição do risco de rescisão – Contrato III	103
9.4	Composição do risco de rescisão – Contrato IV	105
9.5	Composição do risco de rescisão – Contrato V	106
9.6	Composição do risco de rescisão – Contrato VI	107

Lista de Tabelas

5.1 Avaliação da Matriz de Risco	44
6.1 Hiperparâmetros - HL	53
6.2 Hiperparâmetros - <i>Grid Search</i>	53
7.1 Exemplos de descrições de objetos contratuais	65
A.1 Visão Geral da Literatura sobre Empresas e Contratos: Método.	123
A.2 Visão Geral da Literatura sobre Empresas e Contratos: Resultados.	126
B.1 Termos frequentemente associadas a cada tópico.	127
C.1 Lista de CNPJs de serviços essenciais.	130

Capítulo 1

Introdução

A identificação de padrões em conjuntos de dados para classificação de itens em grupos, a previsão de eventos futuros ou o reconhecimento de contexto semântico em estruturas textuais são exemplos de aplicações da Inteligência Artificial (IA) voltadas à otimização de tarefas mecânicas. A partir do avanço na capacidade e performance de processamento, tornou-se possível aplicar métodos estatísticos e técnicas de IA para reconhecer e prever padrões de comportamento em grandes volumes de dados históricos.

Na Administração Pública (AP), o processo de transição de serviços analógicos para plataformas digitais tem gerado grandes volumes de dados, abrangendo tanto os serviços quanto os processos de governança interna de cada organização. Nesse sentido, investir em estratégias apropriadas para armazenamento e tratamento desses dados é essencial não apenas para garantir o cumprimento de normas regulatórias, mas também para possibilitar a instrumentalização e compreensão dos desafios e êxitos das gestões por meio de sua análise [9, Art. 4-5]. A extração de conhecimento desses dados oferece insights valiosos, contribuindo para a melhoria contínua dos processos e serviços públicos.

Diante desse cenário, a obrigatoriedade de disponibilização de dados dos órgãos públicos (órgãos do Executivo, Legislativo, Judiciário e outros entes) foi definida pela Lei Nº 12.527 de 18 de novembro de 2011 — a Lei de Acesso à Informação (LAI) [1]. Como forma de prestação de contas à sociedade, com a intenção de fomentar a participação e controle social, a lei garante o amplo acesso a dados de atividades meio e fim, sua divulgação, disponibilidade, integridade e autenticidade. Esse marco eleva ainda mais o potencial de impacto de

¹Lei de Acesso à Informação - Governo Federal

ferramentas que possam automatizar a análise, a identificação de padrões e a extração de conhecimento de massas de dados.

Ao considerar somente a compra de produtos ou serviços pelos órgãos vinculados a cada um dos 5.570 municípios brasileiros – atividades que, via de regra, necessitam de documentos transacionais –, constata-se que centenas de milhares de contratos são firmados anualmente. Pela lei, a execução de todo contrato público deve ser rigorosamente acompanhada e fiscalizada por um representante da AP designado para a função ou por uma pessoa auxiliar contratada, de modo a garantir seu cumprimento, conformidade e eficiência [10; 8, Art. 117; Art. 67]. Para a gestão, isso implica disponibilidade de pessoal para alocação de ao menos um fiscal por contrato, ou de recursos financeiros para contratação de terceiros que auxiliarão o fiscal em sua função. Na prática, o atendimento a esses requisitos legais pode ser deficiente devido ao alto volume de contratos ou à escassez de pessoal, inviabilizando o acompanhamento e regulação adequados aos respectivos instrumentos.

Ademais, as Controladorias, Ministérios Públicos e Tribunais de Contas representam outro nível de controle para garantir o bom uso dos recursos públicos, atendendo aos interesses da população e assegurando a responsabilização penal ou civil em casos nos quais práticas irregulares são identificadas. Esses órgãos operam de forma tipicamente reativa, conduzindo inspeções em resposta a denúncias ou com base no conhecimento prévio de seus servidores [55]. Devido à baixa escalabilidade desse modo de funcionamento, tanto em relação ao volume crescente de dados quanto à limitação de recursos humanos disponíveis, o uso de estratégias de priorização e seleção de casos para orientação de ações é uma prática comumente adotada.

Levando em consideração o volume de contratos dos municípios e o potencial para automação de atividades repetitivas, a AP representa um cenário promissor para a experimentação e aplicação de técnicas de IA e Aprendizagem de Máquina (AM) na resolução de problemas. Nessa perspectiva, os governos federal, estaduais e municipais têm investido no desenvolvimento de ferramentas que promovam o cumprimento das regulações, melhorem a gestão dos dados gerados e otimizem o uso dos recursos humanos e financeiros disponíveis². Desse modo, os resultados positivos dessas iniciativas incluem a geração de economias por meio da implementação de estratégias para prevenir o descumprimento contratuais, normati-

²Como as robôs Alice, Sofia e Monica ajudam o TCU a caçar irregularidades em licitações.

vos ou o uso indevido de recursos públicos, além da otimização de processos e mitigação dos efeitos da escassez de recursos humanos [21], p. 29]. Ainda, como benefício geral, uma vez que os muitos dos dados gerados por estados e municípios possuem certa correspondência semântica e estrutural, ferramentas e métodos desenvolvidos podem ser adequados ou servir como inspiração para outras instituições interessadas.

A seguir, apresenta-se a motivação deste trabalho, seus objetivos e suas contribuições para os campos da AM e da AP, além da descrição da estrutura deste documento.

1.1 Motivação

O contrato público formaliza um vínculo entre os entes público e privado para fornecimento de bens ou serviços à população, garantindo o bem-estar social. Educação, saúde, assistência social, construção, serviços profissionais, habitação, fornecimento de equipamentos, limpeza e manutenção são exemplos de áreas estratégicas em que os governos nacional e subnacionais desempenham papéis essenciais, recorrendo à contratação de empresas terceiras para suprimento de suas necessidades.

Sob esse viés, a amplitude da responsabilidade governamental em fornecer assistência à população reflete diretamente na abrangência temática dos contratos públicos. Na AP federal, somente no ano de 2023, foram homologados mais de 189 mil processos de compra, totalizando a expressiva cifra de R\$ 171.354.167.865,56 (mais de cento e setenta bilhões de reais)³. De maneira geral, cada compra pode estar vinculada a um ou mais fornecedores e resultar em um ou mais contratos (ou documento equivalente, quando conveniente e possível [10]; [8], Art. 95, Art. 62]).

Entre os princípios estabelecidos pela lei vigente de licitações e contratos (Lei 14.133/2021), destaca-se o dever de contratar com base nos princípios de igualdade e não discriminação [10]. Isso significa garantir que todos os pleiteantes de uma licitação (licitantes) sejam tratados de forma justa e sem privilégios. Assim, frente à impossibilidade, justificada, de contratar com base na reputação ou no desempenho prévio das empresas concorrentes, cabe à AP atuar como gestora dos recursos da sociedade, garantindo seu uso eficaz e eficiente na aquisição de bens e serviços. Nesse contexto, mecanismos de governança são

³Fonte: Painel de Compras, Governo Federal – Processos homologados.

propostos para nortear, monitorar e avaliar a gestão pública.

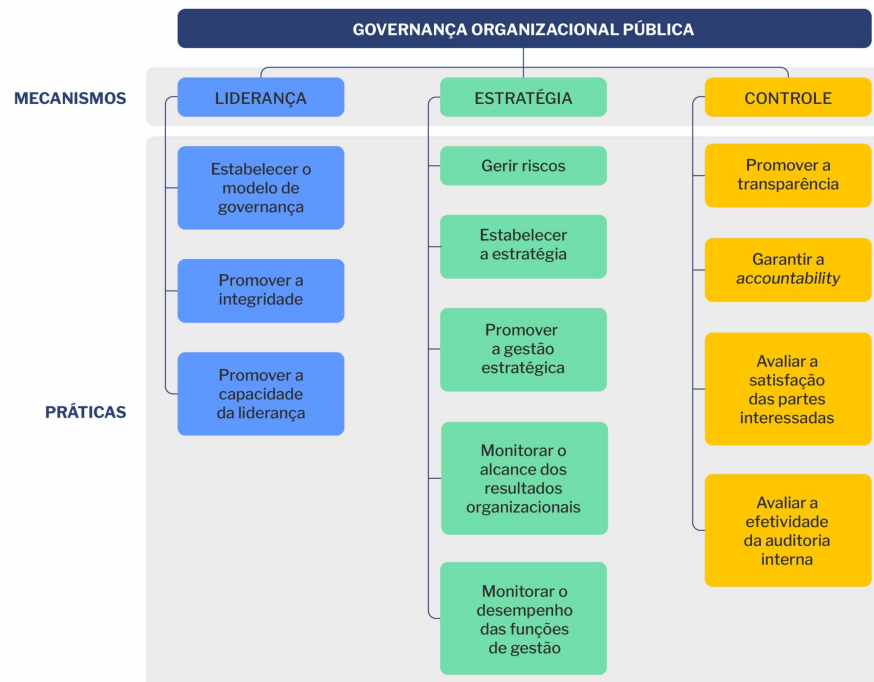


Figura 1.1: Mecanismos e práticas de governança pública organizacional - Fonte: Manual de Licitações e Contratos, Tribunal de Contas da União, 2020, p. 52.

O manual de jurisprudência e orientações sobre contratos e licitações do Tribunal de Contas da União (TCU) explicita mecanismos e práticas para uma governança pública adequada⁴, ilustrados na figura 1.1. Nesse contexto, a mitigação de irregularidades, a gestão de riscos, a garantia de integridade, a eficiência e a transparência são mecanismos fundamentais para o sucesso da gestão contratual [10]. Assim, a definição e o cálculo de riscos associados aos processos internos das instituições são práticas de controle comuns, que tornam tangíveis sua antecipação, monitoramento e mitigação [21, p. 23].

Nesse sentido, com a possibilidade de o contrato estar sujeito a falhas de planejamento, falhas de execução, corrupção envolvendo agentes públicos e/ou privados e outros eventos, cada qual com suas chances de ocorrência em cenários particulares, a própria lei de licitações e contratos define a permanente gestão de riscos e manutenção de controles preventivos [10, Art. 169], com o objetivo de "promover um ambiente íntegro e confiável, assegurar o alinhamento das contratações ao planejamento estratégico e às leis orçamentárias e promover eficiência, efetividade e eficácia em suas contratações" [10, Art. 11]. Com suporte na li-

⁴Licitações e contratos – Orientações e jurisprudência do TCU.

teratura e na prática das instituições e agentes, por meio da estimação e análise dos riscos associados aos contratos públicos, é possível:

1. Garantir um melhor aproveitamento dos recursos públicos e geração de oportunidades com a prestação dos serviços e o fornecimento de bens à população;
2. Auxiliar na orientação de esforços fiscalizatórios em contratações vigentes com base em evidência, a fim de possibilitar um acompanhamento mais rigoroso em contratações possivelmente arriscadas;
3. Economizar recursos humanos e financeiros em investigações reativas sobre contratos previamente problemáticos — uma vez que, frequentemente, é mais oneroso tentar recuperar recursos desviados de seus fins (sem nenhuma garantia de êxito) do que o recurso em si;
4. Aprimorar continuamente processos administrativos, em especial o de contratação, por meio da análise do comportamento prévio de contratados e contratantes;
5. Materializar o conhecimento tácito de atores envolvidos na gestão, controle e fiscalização, promovendo transparência em decisões administrativas e fortalecendo a prestação de contas das instituições.

Casos práticos e a literatura, seja ela proveniente ou não de instituições da AP, apontam para a IA e AM como tecnologias promissoras na resolução de desafios críticos e no auxílio às práticas de governança – em especial na gestão de riscos e na instrumentalização do acompanhamento de gastos. Nesse viés, a previsão do risco de corrupção associado a unidades de gestão públicas [3; 14], a estimação do risco de corrupção associado a empresas da AP federal [47; 48; 52], a identificação de gastos anormais por deputados [29], a detecção de figurantes em pregões eletrônicos ou lances anômalos em pregões [18], bem como avaliação do risco do fracionamento de licitações [12], são exemplos do uso dessas tecnologias.

Diferentes tipificações de risco, granularidade de ocorrência de observações, desenhos experimentais e contextos foram apresentados em trabalhos anteriores que abordam AM aplicada à estimação de risco de entidades na AP. O risco de fraude ou de possíveis irregularidades contratuais é mais comumente abordado na literatura temática, enquanto os desenhos

experimentais consideram comportamentos genéricos ou meras indicações de risco como rótulos de condutas históricas.

No contexto brasileiro, o risco de falha em contratos federais é materializado por meio do uso de sanções prévias sofridas pelas empresas responsáveis pelos contratos [32; 48; 52]. Além disso, o risco de fracionamento de licitações e contratos federais é abordado com base na opinião de especialistas para validação do modelo proposto [12]. Ainda, comportamentos anômalos em compras federais são modelados e reconstruídos a partir de uma abordagem inovadora à época [26].

Na Colômbia, dados dos governos federal e municipais são utilizados para previsão do risco de possíveis condutas ilegais ou indesejadas em contratos, considerando o comportamento prévio das empresas e um indicador de risco como rótulo para definição de classes [27]. Em uma abordagem inovadora, em parceria com o Banco Mundial, a modelagem de comportamentos de risco de fraude, corrupção e conluio foi proposta utilizando contratos de abrangência internacional a partir de rótulos tanto positivos (contratos sabidamente irregulares) quanto negativos (contratos não irregulares) [30].

Além disso, a modelagem dos comportamentos de risco nos exemplos citados, por exemplo, exige não apenas domínio das técnicas de AM, mas também um entendimento profundo do complexo processo de contratação pública. Por outro lado, a colaboração com entidades da AP, órgãos de controle e sociedade civil deixa evidente que, apesar de oportunas, as soluções utilizando AM ainda não são adotadas de maneira sistêmica, sendo aplicadas majoritariamente a desafios pontuais. Ademais, a insegurança no uso das ferramentas, incertezas sobre o seu modo de funcionamento e possíveis impactos negativos no trabalho realizado são possíveis gargalos para a adoção plena deste tipo de tecnologia.

Estudos anteriores, em geral, exploram a viabilidade técnica de seu constructo em relação às técnicas adotadas e aos dados disponíveis. Especificamente sobre o tema da modelagem do risco de contratos, existem lacunas quanto ao desenho experimental, à falta de detalhes acerca da engenharia de características, à ausência de explicações acerca da composição de risco para as previsões feitas pelo modelo a partir dessas características. Ainda, é pertinente destacar que a literatura pública atual não reflete completamente os esforços de praticantes (pertencentes ou não ao setor público) no tema, uma vez que muitos trabalhos são de consumo interno dos órgãos ou apenas não foram publicados em periódicos, revistas ou si-

milares.

A motivação para a escolha deste tema de pesquisa esteve principalmente ligada às lacunas relacionadas ao desenho da modelagem de risco contratual, à quantificação experimental do impacto de conjuntos de características na capacidade preditiva do modelo e à quantificação experimental da eficácia de um método de estimação de risco utilizado na prática. A partir disso, espera-se trazer maior clareza e refinamento na aplicação de métodos de AM na estimação de risco, consolidando práticas já existentes e explorando novas possibilidades. Além disso, busca-se aprofundar a compreensão da natureza e composição do comportamento de risco de rescisão como uma característica tangível.

Assim, utilizando mutações contratuais (paralisações, rescisões, suspensões, impedimentos e sustações) como rótulos de risco, este trabalho, no contexto municipal paraibano, avalia experimentalmente um método em uso para estimação de risco; detalha e avalia a modelagem de risco a partir da perspectiva do risco de rescisão a nível de contrato; experimenta estratégias para suplantear problemáticas comuns, como o desbalanço entre classes e indisponibilidade de dados; e, por fim, propõe e analisa o impacto de características na composição do risco de rescisão dos contratos. Com isso, espera-se que os esforços aqui refletidos contribuam para expandir o conhecimento na aplicação de técnicas de AM em ferramentas de controle, na tarefa de modelagem do risco de contratos públicos e no aprimoramento do processo de contratação.

1.2 Objetivos

Tendo em vista a necessidade de priorização de esforços de fiscalização ou investigação no montante de contratos produzido por cada município paraibano, o objetivo deste trabalho é aprimorar a modelagem de risco de rescisão em contratos públicos e aprofundar a compreensão do comportamento de risco através da observação de tendências. Para isso, foram avaliadas experimentalmente mudanças no desenho da modelagem e nas características utilizadas como insumo no processo de aprendizagem, além de analisada, de forma empírica, a influência dessas variáveis na previsão de risco do modelo.

1.3 Contribuições

A quantificação da influência do desenho da modelagem na eficiência preditiva do classificador de risco e a aplicação de técnicas do estado da arte no contexto de contratação pública municipal são contribuições relevantes para a aprendizagem de máquina aplicada ao setor público. No âmbito da administração pública, proporciona uma compreensão mais aprofundada sobre o efeito de cada característica específica na composição do risco de rescisão contratual, além de contribuir para a discussão da validade do ferramental quanto ao seu uso prático. Dessa forma, espera-se que este trabalho possa servir de ponto de partida para o desenvolvimento e implementação de dispositivos de controle ativos baseados em AM para contratações municipais mais efetivas e eficientes.

Em específico, realiza-se a avaliação experimental de uma ferramenta utilizada para estimar o risco de empresas, ampliada para contratos públicos firmados por essas instituições. Com isso, quantifica-se empiricamente seu impacto prático e estabelece-se uma base comparativa para os modelos seguintes.

Foram avaliadas, experimentalmente, estratégias para o refinamento da modelagem do comportamento de risco de rescisão, através do uso de um conjunto de rótulos inédito, estratégias para mitigação do desbalanceamento entre classes e o efeito de características preditoras utilizadas previamente ou propostas neste trabalho.

Por fim, analisou-se como cada característica utilizada como preditor no processo de aprendizagem do modelo impacta, na prática, a composição do risco de rescisão dos contratos de maneira geral e especificamente para contratos corretamente classificados. Essa análise fortalece a confiança no modelo ao proporcionar maior clareza no entendimento do risco, compara as previsões com o senso comum de especialistas, e pode auxiliar outros praticantes na produção de novas características, na adoção da ferramenta por órgãos de controle ou no aperfeiçoamento das características já utilizadas ou, ainda, no aprimoramento da governança dos contratos públicos.

1.4 Estrutura do Documento

O restante deste documento está estruturado da seguinte maneira:

Capítulo 2 expõe definições teóricas necessárias a este trabalho;

Capítulo 3 examina trabalhos anteriores com proposta semelhante, destacando pontos fortes e lacunas desses estudos;

Capítulo 4 apresenta os materiais e metodologias utilizadas nesta pesquisa;

Capítulo 5 avalia experimentalmente o estado da prática na estimação de risco na Paraíba;

Capítulo 6 explora a estimação de risco em nível de contrato, suas oportunidades e desafios;

Capítulo 7 explora o uso ou impacto da ausência de conjuntos de características na eficácia da modelagem de risco;

Capítulo 8 explica detalhadamente o tipo de impacto que cada característica possui no comportamento de risco de rescisão;

Capítulo 9 apura os desdobramentos reais de contratos corretamente classificados pelo modelo;

Capítulo 10 apresenta as conclusões deste estudo e oportunidades de trabalhos futuros.

Capítulo 2

Fundamentação Teórica

Este capítulo apresenta os fundamentos teóricos que embasam o processo de elaboração e percurso deste estudo, necessários para sua compreensão.

2.1 Contratação pública

O processo de contratação pública é compreendido por etapas, que nesta seção serão resumidamente descritas. A primeira etapa, destinada à licitação, visa especificar e tornar pública a necessidade de determinado bem e/ou serviço e selecionar a proposta mais vantajosa para seu provimento, de modo a promover o desenvolvimento nacional e garantir um tratamento igualitário perante a lei (princípio da isonomia) para com as pessoas físicas ou jurídicas participantes (licitantes) [1]. Essa etapa contempla as fases de preparação e planejamento do edital de licitação, divulgação do edital, recebimento de propostas e lances (quando oportuno), julgamento e habilitação das propostas, apresentação de recursos e homologação dos licitantes vencedores [10, Art. 17]. Por outro lado, os participantes do processo estão responsáveis pela ciência do documento do edital, apresentação de propostas, desistência do certame, apresentação de recursos e impugnações [44].

A segunda etapa, a contratação, estabelece as condições para o fornecimento do bem ou execução da prestação do serviço e responsabilidades das partes [10, Art. 89 §2º], cabendo aos licitantes adjudicatados⁵ (pessoas físicas ou jurídicas) firmarem ou não o termo de contrato

⁵Adjudicação - ato de entrega do objeto licitado ao licitante vencedor e convocação do mesmo para a assinatura do contrato.

ou documento equivalente. Com a finalidade de formação de um vínculo, o documento do contrato estabelece, entre outras determinações, o objeto e suas características, como preço, prazo e referência ao edital de licitação originário ou ato para contratação direta [1] [10], Art. 92]. Apesar de obrigatório, o contrato pode ser substituído por instrumento correspondente quando conveniente [10], Art. 95].

Já a reserva orçamentária do gasto estimado na licitação é definido posteriormente em contrato é denotada pela fase de empenho. A fase de liquidação caracteriza-se pela conferência da prestação do serviço ou recebimento do bem adquirido. O pagamento é a transferência efetiva do montante acordado em contrato ao fornecedor. Essas três fases compõem a execução orçamentária, etapa transversal à contratação pública e posterior ao planejamento orçamentário anual.

A nova Lei de Licitações e Contratos (NLLC), Lei 14.133 de 1º de abril de 2021⁶, regimenta o Artigo 37 da Constituição Federal Brasileira e "estabelece normas gerais sobre licitações e contratos administrativos [...] no âmbito dos Poderes da União, dos Estados, do Distrito Federal e dos Municípios". Ainda, são também detalhadas em lei as modalidades, fases e tipos de licitação, tipos de contratos e aspectos da qualificação das empresas concorrentes, possíveis sanções administrativas ou penais para toda e qualquer prática que obstrua as diretrizes estabelecidas [40].

Além disso, é conveniente citar a antecessora da NLLC, a Lei 8.666, de 21 de junho de 1993⁷, antiga Lei Geral de Licitações e Contratos (ALGL). Tanto a ALGL, quanto a Lei do Pregão (10.520, de 17 de julho de 2002) e parte da Lei do Regime Diferenciado de Contratações (12.462, de 4 de agosto de 2011) foram revogadas a partir da vigência da NLLC [10], Art. 191, 193].

Desse modo, fundamentadas na NLLC, para fins de atualização, as explicações desta seção e do Capítulo I são pontos comuns entre ambas as leis e não visam expor exaustivamente suas minúcias, mas sim sintetizar operações quanto à contratação pública e possibilitar a compreensão deste trabalho. Assim, vigente à época de assinatura e execução do conjunto de contratos objeto deste estudo, a ALGL servirá como referência ao restante do texto.

⁶ Lei 14.133/2021 - nova Lei de Licitações e Contratos

⁷ Lei 8.666/1993 - antiga Lei Geral de Licitações e Contratos

2.1.1 Contrato

Contratos públicos são instrumentos pelos quais órgãos da administração direta, autarquias, fundos especiais, fundações, empresas públicas e demais entidades controladas, direta ou indiretamente, pela União, estados ou municípios realizam a aquisição de bens e serviços por meio da contratação de empresas [8, Art. 1] [20]. Nesse contexto, fundamentais para suprir as necessidades da população e governo, esses instrumentos são, de modo geral, regulados pela lei de licitações e contratos em vigor, por suas próprias cláusulas, além de seguirem os preceitos do direito público [8, Art. 54].

No entanto, durante a execução contratual, não é incomum que a ineficiência na gestão, falhas, irregularidades ou outros incidentes findem o vínculo firmado antes da sua conclusão. Se por um lado é de interesse público o manutenção do contrato e finalização da prestação do serviço, também é importante que a fiscalização desses contratos ocorra sem onerar em excesso sua realização. À vista disso, a própria gestão responsável, sociedade civil e órgãos públicos de controle têm um papel importante no acompanhamento desses contratos e dos recursos investidos neles, garantindo que o objetivo da contratação seja alcançado.

Ademais, na intenção de resguardar o interesse público e a própria gestão, assim como no intuito de comprovar o cumprimento dos resultados acordados, o artigo 67 da ALGL estabelece que "a execução do contrato deverá ser acompanhada e fiscalizada por um representante da Administração especialmente designado, permitida a contratação de terceiros para assisti-lo e subsidiá-lo de informações pertinentes a essa atribuição". Na prática, alocar fiscais suficientes para cada contrato firmado por entes nacionais e subnacionais é um obstáculo ao cumprimento da lei, tendo em vista a escassez de recursos humanos e/ou financeiros. Como consequência, esse desafio, por vezes, distancia entidades e órgãos governamentais do êxito de seus objetivos para a manutenção do seu pleno funcionamento.

2.1.2 Desfazimento e mutação contratual

A execução parcial ou inexecução total são razões válidas para a extinção de um contrato. A lentidão, o não cumprimento ou cumprimento irregular de cláusulas, especificações ou prazos, além de irregularidades, alteração social ou estrutural na empresa que restrinja a capacidade de execução ou até mesmo a falência da empresa são motivos que podem justi-

ficar a quebra de vínculo. Nessa perspectiva, as hipóteses de desfazimento contratual, tanto da parte contratante quanto contratada, estão especificadas na seção V, *Da Inexecução e da Rescisão dos Contratos*, da ALGL. Para referência, na NLLC, essas hipóteses estão previstas no Capítulo VIII, *Das Hipóteses de Extinção dos Contratos*. Nesse caso, as obrigações entre as partes do contrato encerram-se.

Já alterações (mutações) contratuais podem ocorrer para incluir mudanças nas condições inicialmente definidas entre as partes do contrato ou, simplesmente, para alterar seu estado. A seção III, artigo 65 da ALGL, *Da Alteração dos Contratos*, rege alterações no conteúdo do contrato, que podem depender ou não de análise jurídica prévia. Aqui, o contrato vigente permanece válido, mas seus termos são ajustados.

Assim, um dos interesses deste trabalho é estimar as chances de que motivos relativos à quebra do contrato ou alterações que comprometam o alcance de seu propósito ocorram com base no histórico da empresa responsável ou das próprias características do contrato. Os motivos de mutação ou desfazimento contratual nos contratos municipais paraibanos são descritos como: Aditivo, Apostilamento, Retomada, Ordem de Serviço, Paralisação, Rescisão, Suspensão, Impedimento ou Sustação. É necessário pontuar que a quebra ou mutação contratual não é, necessariamente, ocasionada por irregularidade, visto que pode ocorrer por necessidade da AP ou por acordo entre as partes.

Detalhes sobre os motivos de mutação e quebra contratual úteis a este trabalho são descritos na seção 4.1 deste documento.

2.1.3 SAGRES

No Estado da Paraíba, o SAGRES⁸ (Sistema de Acompanhamento da Gestão dos Recursos da Sociedade.) é o sistema responsável por receber, armazenar e disponibilizar dados sobre o processo de contratação e execução da despesa pública: licitações, contratos, empenhos, liquidações, pagamentos, dados sobre pessoas físicas e jurídicas envolvidas nesses processos, além de outras informações. O sistema foi desenvolvido pelo Tribunal de Contas do Estado (TCE) da Paraíba, no ano de 2002, como uma forma de atender aos requisitos da Lei de Responsabilidade Fiscal e é alimentado automaticamente pelos órgãos sob a jurisdição deste TCE [49].

⁸SAGRES - Tribunal de Contas do Estado da Paraíba.

O SAGRES conta com diversos módulos. Este estudo utiliza o módulo SAGRES Municipal que conta com cerca de 190 tabelas e é responsável por armazenar dados de todas as unidades gestoras dos 223 municípios paraibanos. A ideia e a arquitetura do SAGRES foram replicadas e aperfeiçoadas por diversos estados brasileiros, sendo uma prática de transparência pública notável. Nesse contexto, o SAGRES representa, para a esfera municipal, o que o Portal de Transparência ⁹ representa para a esfera federal do Governo.

No Capítulo 6, serão exibidos mais detalhes sobre como os dados disponíveis no SAGRES Municipal são utilizados para compor as características que servem como variáveis preditoras na modelagem.

2.2 Aprendizagem de máquina

A AM é uma subárea da IA que se propõe a estudar e desenvolver algoritmos capazes de aprimorar seu resultado com base em dados históricos. Ao modelar determinado comportamento, considera-se um conjunto de dados X_{ij} formado por j preditores (ou características de entrada) com $j = \{1, 2, \dots, p\}$ para cada observação i , com $i = \{1, 2, 3, \dots, n\}$, e Y_i representando a variável resposta da i -ésima observação. Assim, esse conjunto de dados pode ser representado por $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, onde $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})^T$ [34, p. 21], exibido em 2.1. O objetivo da aprendizagem é, portanto, associar o conjunto de observações e suas respectivas características à variável resposta. Isto é, espera-se encontrar uma função f tal que $Y = f(X)$ para cada par (X, Y) .

$$f \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix} \rightarrow \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \quad (2.1)$$

Hastie *at al.* [31] definem que problemas resolvidos por meio da aprendizagem de máquina podem ser divididos em duas grandes categorias: aprendizagem supervisionada e não supervisionada.

⁹Portal da Transparência - Governo Federal.

Aprendizagem supervisionada

Na aprendizagem supervisionada, para cada observação x_i (e respectivos preditores) haverá uma resposta y_i associada. Aqui, o objetivo é encontrar um modelo matemático que relacione a variável resposta (ou variável dependente) às características preditoras (ou variáveis independentes), considerando todas as observações. Com isso, é possível compreender a relação entre características preditoras e respectivas respostas em dados já conhecidos, ou prever uma resposta para observações futuras [34, p. 26].

Sob essa perspectiva, utiliza-se uma parte das observações do conjunto de dados para treinamento do modelo, ou, em outras palavras, para ajustar a função nas características preditoras em relação à variável resposta — essa parcela dos dados é chamada de conjunto de treino. Enquanto o restante dos dados, denominado conjunto de teste, é utilizado para avaliar a eficácia do modelo e sua capacidade de generalização de previsão para dados não conhecidos. A avaliação é realizada utilizando os preditores do conjunto de testes como entrada do modelo e comparando sua previsão com os valores da variável resposta já sabida para cada observação.

Grosso modo, modelos cuja variável resposta é quantitativa são denominados modelos de *regressão*, enquanto aqueles que preveem uma variável resposta qualitativa (ou categórica) são denominados modelos de *classificação*. Um bom modelo não é, necessariamente e sempre, o que atinge valores mínimos de erro em sua previsão, mas também o que representa um preditor útil ao contexto para o qual foi projetado e modelado.

Neste trabalho, aborda-se a estimação do risco rescisão de contratos públicos municipais a partir da aprendizagem supervisionada, utilizando como variável resposta contratos sabidamente problemáticos no passado.

Aprendizagem não supervisionada

Já a aprendizagem não supervisionada carece de resposta associada a cada observação x_i , não sendo possível definir uma função f , uma vez que não há valor a ser previsto. Aqui, o objetivo passa a ser exploratório, buscando encontrar padrões (de similaridade ou dissimilaridade) entre características ou entre observações que possam ser úteis no uso dos dados. O agrupamento (ou clusterização) é um exemplo de técnica de aprendizado não supervisionado

e baseia-se na junção de observações em grupos, com base na similaridade de características. Nesse tipo de aprendizado, não é trivial mensurar a qualidade dos comportamentos ou agrupamentos produzidos por um modelo, na medida em que não há resultados a serem comparados. Logo, usualmente, recorre-se a heurísticas ou conhecimento especializado do próprio domínio na tentativa de justificar ou avaliar as inferências feitas.

2.3 Desbalanceamento entre classes

A identificação de comportamentos atípicos em cenários financeiros, por meio da aprendizagem de máquina, tem sido uma tarefa amplamente estudada nos últimos anos, por exemplo: fraudes em leilões [2; 18], serviços bancários e seguros, fraude ou corrupção em contratos públicos [48] ou privados [30]. Em cenários onde algoritmos de aprendizagem supervisionada sejam utilizados, é necessário que o conjunto de dados de treinamento contenha um número equivalente de exemplares de cada classe. Dessa maneira, o modelo/função pode ajustar-se e realizar previsões a partir do comportamento de ambas as classes de maneira mais acurada. Entretanto, nesses cenários, uma característica comum dos conjuntos de dados de treinamento dos modelos é o desbalanço, por vezes severo, entre as classes da variável resposta – onde uma possui um tamanho significativamente superior ao outro.

Um modelo treinado em um conjunto desbalanceado terá viés em relação à classe majoritária pela métrica de avaliação, já que as regras de aprendizagem que preveem corretamente um maior número de observações serão relacionadas a essa classe, tendendo a classificar erroneamente observações da classe minoritária. Assim, o uso de estratégias para lidar com o desbalanço entre classes da variável dependente impacta diretamente nos resultados dos modelos de classificação, podendo, ainda, prejudicar sua capacidade preditiva. Em alguns cenários, o uso de técnicas de balanceamento artificial para tornar o tamanho das classes equivalentes no conjunto de dados de treinamento tem demonstrado melhora na performance de modelos preditivos [2].

De modo geral, essas técnicas podem aumentar o número de observações da classe minoritária (*data oversampling* – sobreamostragem em tradução livre do inglês), diminuir o número de observações da classe majoritária (*data undersampling* – subamostragem de dados em tradução livre do inglês) ou combinar aumento e diminuição. No primeiro caso, a

sobreamostragem pode ser realizada pela simples replicação de amostras de dados da classe minoritária ou pela criação de observações sintéticas [16]. No segundo caso, a subamostragem pode ser realizada por meio da seleção de amostras menores da classe majoritária.

A literatura apresenta que métodos para sobreamostragem têm eficácia limitada em conjuntos de dados reais. Por um lado, a replicação aleatória dos mesmos valores e características não produz novas informações ao espaço de características e consequentemente no aprendizado de regras pelo modelo; e observações sintéticas podem comprometer a representatividade dos registros [54]. Por outro, reduzir o volume de observações através da subamostragem acarreta a perda de uma parcela de dados de treino devido ao balanceamento. Uma estratégia alternativa é utilizar o aprendizado sensível ao custo para lidar com o desbalanceamento, atribuindo pesos de importância a cada classe ou observação para ajustar sua importância durante o treinamento, assumindo custos maiores na classificação errônea de amostras da classe minoritária. Diferentes estratégias podem ser úteis, a depender do domínio específico.

2.4 Processamento de linguagem natural

Até então, as técnicas de AM apresentadas focam na modelagem, análise e compreensão de dados numéricos. O Processamento de Linguagem Natural (PLN), subárea da IA, oferece estratégias para resolução de tarefas, a partir de documentos textuais. Essas estratégias consideram a distribuição e relevância de palavras em documentos distintos, a estrutura do corpo de documentos e os aspectos sintáticos e semânticos [5].

2.4.1 Modelagem de Tópicos

A modelagem de tópicos (ou identificação de tópicos) consiste em um tipo de modelagem estatística não supervisionada que faz uso da linguagem, para descoberta de estruturas em uma coleção de documentos textuais (*corpus*). O objetivo dessa técnica é sumarizar estruturas latentes em grandes volumes de documentos, cuja rotulagem manual torna-se inviável [4]. Essa técnica é comumente utilizada em tarefas de redução de dimensionalidade, aprendizado não supervisionado e rotulagem de classes a partir de tópicos.

A estratégia *Online* [33], da *Latent Dirichlet Allocation* (LDA), promete lidar de maneira

eficiente com grandes volumes de documentos e diversos tópicos. A partir de um *corpus*, um conjunto de palavras e um número de tópicos previamente informado como parâmetro, o modelo retorna a distribuição de palavras e tópicos para cada documento. De maneira simples, cada documento pode ser visto como uma mistura de tópicos (em que cada tópico tem um peso específico) e cada tópico como uma mistura de palavras. Já a *Latent Semantic Analysis*(LSA) e a *Probabilistic Latent Semantic Analysis*(pLSA) são outras técnicas comumente utilizadas para identificação de tópicos em documentos textuais.

2.5 XGBoost

O algoritmo de AM *eXtreme Gradient Boosting* (XGBoost) foi utilizado para modelagem do risco de rescisão de contratos públicos. O XGBoost foi utilizado em trabalhos anteriores e tipicamente alcança performances do estado da arte em dados tabulares [51] que possuem escalas similares a do conjunto de dados utilizados neste estudo.

Anterior ao treinamento dos modelos com o XGBoost, é necessário compreender e definir valores dos hiperparâmetros, que podem impactar positivamente o desempenho dos classificadores. Esses parâmetros determinam, por exemplo, a velocidade, a profundidade e o nível de conservadorismo no processo de aprendizado. São eles:

n_estimators (número de estimadores) – representa a quantidade de árvores consideradas na construção do classificador, isto é: o número de rodadas de aprendizagem – $[0, +\infty)$. Um aumento no número de árvores pode melhorar a performance do modelo, mas também pode causar sobreajuste do modelo nos dados ou maior tempo de treinamento.

learning_rate (taxa de aprendizado) – define o ritmo de aprendizado, determinando a contribuição de cada estimador no resultado final – $[0, 1]$. Um valor demasiado baixo retarda o aprendizado e pode requerer uma maior quantidade de estimadores, apesar de possível melhor generalização de aprendizado no conjunto de dados. Uma alta taxa pode levar a um treinamento mais rápido, porém possivelmente com uma cobertura apenas parcial dos padrões disponíveis nos dados.

scale_pos_weight – balanceia o aprendizado de ambas as classes atribuindo-lhes pesos. Tipicamente, esse hiperparâmetro assume o valor da proporção da classe negativa em

relação a classe positiva, como exibido na fórmula [2.2](#)

$$\text{scale_pos_weight} = \frac{\text{Num. Obs. Classe Negativa}}{\text{Num. Obs. Classe Positiva}} \quad (2.2)$$

reg_lambda – este hiperparâmetro é o responsável pela regularização L2 nos pesos das folhas. Quanto maior o valor, mais conservador o modelo.

min_child_weight – define a soma mínima de pesos de todas as observações em cada nó filho do estimador. Quanto maior o valor deste hiperparâmetro, mais conservador será o modelo, entretanto valores muito altos podem levar a sobreajuste.

Além dos hiperparâmetros descritos acima, outros podem ser experimentados para refinamento do classificador gerado [10](#).

2.6 Explicabilidade em modelos de aprendizagem

Com o avanço tecnológico, a expectativa de geração de valor é proporcional à disponibilidade e volume de dados coletados pelos setores público e privado. A complexidade de modelos de aprendizagem de máquina é também reflexo da complexidade dos padrões presentes nos dados de treinamento. Modelos lineares, como uma regressão, por exemplo, irão identificar relações lineares entre as características preditivas no aprendizado, podendo o relacionamento entre as características ser entendido através de seus coeficientes. Analogamente, modelos não-lineares poderão identificar relacionamentos menos triviais ao custo de serem baseados em estruturas mais complexas (camadas de interação na aprendizagem, funções de ativação, parametrização extensa, por exemplo).

A interpretabilidade de um modelo de AM pode ser definida como o nível no qual um ser humano consegue entender uma decisão a partir de uma causa [\[38\]](#). A aplicação de modelos de aprendizagem de máquina em áreas críticas, como a medicina, justiça e finanças,

¹⁰ Documentação XGBoost - Parâmetros e hiperparâmetros

por exemplo, tornou ainda mais fundamental que seres humanos possam interpretar e explicar o comportamento dos modelos criados, principalmente para validação de sua eficácia e constructo, descobrimento de conhecimento, identificação de possíveis vieses ou mesmo para a justificativa de seu comportamento preditivo [39]. Além disso, a interpretação das previsões de um modelo pode ser comunicada aos especialistas na área de interesse e, a partir disso, podem ser realizadas melhorias no conjunto de dados utilizado no aprendizado ou no algoritmo de treinamento [56].

Diante disso, o conceito de interpretabilidade será utilizado neste estudo para tornar possível a compreensão das classificações de risco de rescisão contratual realizadas pelo modelo, a partir das características utilizadas no aprendizado.

2.6.1 Valores SHAPley

A literatura conta com uma diversidade de métodos para interpretação, que podem ser específicos para uma classe de modelos (e, nesse caso, também são considerados os modelos intrinsecamente interpretáveis, como os lineares), ou genéricos suficientes para abarcar modelos de AM no geral. Nesse sentido, métodos locais explicam previsões individuais para cada observação, enquanto métodos globais fornecem uma explicação sobre o comportamento do modelo como um todo. Neste trabalho, serão utilizados os valores SHAPley, um método local e genérico para interpretação das previsões de um modelo de risco de rescisão de contratos.

Sucintamente, valores SHAPley são utilizados para determinar quanto e de que modo o valor de cada característica contribuiu para o valor de previsão de uma observação específica em relação à previsão média do modelo para todo o conjunto de dados. Em teoria, o cálculo do valor SHAPley é a contribuição marginal média da presença ou ausência do valor da característica dentre todas as possíveis combinações de características. Entretanto, na prática, a seleção de amostras dentre as combinações possíveis para estimar a contribuição é uma aproximação que possibilita o equilíbrio entre a conclusão da tarefa em tempo hábil razoável, uma vez que, quanto maior o número de combinações, maior será o tempo necessário para computação [39]. Por fim, a soma dos valores SHAPley sempre será igual ao valor médio das previsões por um modelo treinado utilizando todo o conjunto de características.

Desse modo, a possibilidade de explicações contrastantes e explicações individuais são

propriedades que motivaram o uso dos valores SHAPley para interpretação das previsões. Nesse cenário, é possível captar contribuições não lineares de uma mesma característica, inferir comportamentos gerais através do conjunto de comportamentos individuais e descrever comportamentos de risco de subconjuntos de contratos específicos. Assim, o SHAP [36] é um método otimizado para computação dos valores SHAPley, implementado e disponibilizado em diversas linguagens de programação, e será utilizado neste trabalho.

Capítulo 3

Referências e estado da arte

Neste capítulo são descritos trabalhos prévios que abordaram, direta ou indiretamente, os temas de estimação de risco ou detecção de comportamentos atípicos em entidades da AP.

3.1 Estimação de Risco

A combinação de mineração de dados e AM tem se tornado um conjunto de ferramentas populares na detecção e previsão de fraude ou comportamentos anômalos no âmbito financeiro [57]. Administrações de órgãos públicos e privados têm investido nessa área, concentrando esforços no processamento e na análise de dados cadastrais e comportamentais de suas principais entidades para gerar inteligência. Nesse contexto, a criação de indicadores de risco ou regras baseadas em comportamentos suspeitos, bem como a detecção de padrões de comportamento e similaridades, são exemplos de tarefas relevantes, listadas a seguir.

3.1.1 Empresas Licitantes

A empresa que se submete ao processo de contratação pública (licitante) é a principal entidade modelada no domínio da estimação de risco. Algoritmos de regressão logística e árvore de decisão foram utilizados para modelar características sobre 1.000 empresas contratadas e classificá-las quanto às suas chances de serem sancionadas (utilizando rótulos provenientes do Cadastro de Empresas Inidôneas e Suspensas – CEIS) [47]. Adicionalmente, as classes de risco dos dados de treinamento foram artificialmente balanceadas por meio da amostragem

aleatória da classe de maior tamanho.

A acurácia geral dos modelos de regressão logística (62,5%) e árvore de decisão (63%) mostrou-se promissora à época, ao considerar que esse foi um dos primeiros esforços da aplicação da AM para estimação de risco no país. Esse trabalho foi realizado por integrantes do Observatório da Despesa Pública, uma unidade interna da Controladoria Geral da União (CGU). Sua principal contribuição está na avaliação de características relevantes na distinção entre fornecedores possivelmente inadimplentes, destacando atributos relativos ao tamanho ou porte da empresa como bons discriminantes.

Ainda no âmbito federal, Sales *et al.* [48], componentes da CGU, propuseram o uso de dados sobre contratações e compras governamentais provenientes do Sistema Integrado de Administração de Serviços Gerais (SIASG), para medir o risco associado a contratos públicos. O histórico do CEIS foi novamente utilizado como rótulo para contratos de empresas de alto e baixo risco. *Naïve Bayes*, *Tree Augmented Naïve Bayes*, *Tabu Search* e *Hill-Climbing* foram os modelos de classificação avaliados, não havendo diferença estatisticamente significativa entre suas performances. Apesar disso, o primeiro modelo foi selecionado para o cálculo de métricas avaliativas no conjunto de testes, seu intervalo de confiança com 95% para acurácia variou entre 0,69 e 0,77, acurácia = 73% e sensibilidade (TPR) = 76%. Apesar de propor estimar o risco em contratos públicos, o grão das observações utilizadas para modelagem e rótulo são empresas (rotuladas como "Good" e "Bad", "boas" e "ruins" em tradução livre do inglês), sendo estimado, ao final, o risco de cada uma. Também é válido destacar que as características preditoras listadas são descritores sobre empresas.

No âmbito estadual, [25] *et al.* descrevem e normalizam as características utilizadas como preditores de risco. Além disso, em um cenário pouco desbalanceado, compara o uso do balanceamento artificial de classes com sua ausência. Nesse trabalho, o rótulo das observações é definido a partir da citação da empresa em relatórios de auditoria por envolvimento em fraudes ou rescisões contratuais. O algoritmo J48 (árvore de decisão) demonstrou melhor performance na classificação de risco das empresas.

3.1.2 Contratos e Compras

Em compras do Governo Federal, gastos sobre TI foram utilizados como entrada do algoritmo *Deep Auto-Encoder* para identificação de despesas anômalas [26]. Dentre os 12

modelos avaliados, o de melhor desempenho era composto por 15 camadas escondidas, 5 iterações, função de ativação *Tahn*, e seu Erro Quadrático Médio (*Mean Squared Error* - MSE) ficou em torno de 0,001. O modelo de melhor desempenho foi aplicado nos dados de teste e foi avaliado o erro na reconstrução de cada registro. A estratégia de validação utilizada foi selecionar os registros com maior MSE na sua reconstrução e verificar suas características. A partir disso, seria possível priorizar as ocorrências onde o erro foi maior em investigações futuras.

Os algoritmos KNN, Regressão Logística, Análise Discriminante, SVM e Floresta Aleatória modelaram a previsão do risco de falhas em contratos públicos federais, a partir de dados do ComprasNet, SICAF, RAIS, RF, TSE [32]. Novamente, os rótulos para as classes de risco foram extraídos do CEIS. O modelo utilizando KNN obteve uma performance superior aos demais, com 93,5% de acurácia.

A primeira aplicação de redes neurais profundas no contexto da auditoria governamental comparara o desempenho desse algoritmo com os de rede neural tradicional, regressão logística e análise da função discriminante [52]. Em parceria com a CGU, foram modeladas as características de contratos acima de 1 (um) milhão de reais e suas respectivas empresas responsáveis. Sete fontes de dados foram utilizadas para geração de características, e dados sobre sanções prévias das empresas responsáveis (provenientes do CEIS) serviram como rótulo das classes de risco. No conjunto de treinamento, as classes de risco foram artificialmente balanceadas com sobreamostragem. O modelo utilizando redes neurais profundas demonstrou melhores resultados, com acurácia = 97%; especificidade (TNR) = 94%; precisão = 0,4; AUC = 0,87 e $F_{0.5} = 0,56$. A pouca capacidade interpretativa do aprendizado e decisões do modelo é um desafio no uso dessa classe de algoritmos.

No contexto internacional, Grace *et al.* [30] avaliaram a eficácia de modelos utilizando algoritmos de floresta aleatória, AdaBoost, classificação com vetores de suporte (SVC), *Gradient Boosting Machine* (GBM) e KNN para previsão de risco de fraude, corrupção e conluio em contratos internacionais. Nesse estudo desenvolvido em parceria com o Banco Mundial, foram usadas características relativas a empresas e contratos, e dados sobre investigações em contratos passados como rótulo das classes de risco. Uma característica crucial dessa investigação é que foram propriamente avaliados e rotulados tanto contratos que levaram a algum indício de irregularidade, como também contratos onde nada foi encontrado. Diferen-

temente de outros estudos, a parcela de observações de baixo risco não apenas não possui a característica de interesse, como também, de fato, possui a característica de não-interesse. Em suma, contratos rotulados como baixo risco não apenas não são de alto risco, como são comprovadamente de baixo risco — uma base que contém, de fato, ambas as classes de risco. A partir da precisão (0,7) como principal métrica de avaliação, o modelo utilizando GBM surge como o de melhor resultado.

A partir da modelagem de contratos da Colômbia (Sistema Eletrônico de Contratação Pública – SECOP), de 2011 a 2015, o modelo utilizando GBM atingiu a melhor performance, com um IC para AUC-ROC que variou entre 0,78 e 0,91. Dados sobre risco de irregularidades ou potenciais violações foram utilizados como rótulos de risco. Por utilizar observações e rótulos para as observações no nível/grão do contrato municipal, esse trabalho, desenvolvido por Gallego *et al.* [27], é o mais próximo do proposto por esta pesquisa. Contudo, dois dos três rótulos experimentados estiveram no nível de empresa, enquanto o terceiro não representava falhas contratuais de fato, mas mera indicação de risco.

No contexto espanhol, foi apresentada uma análise descritiva quantitativa sobre 58.337 contratos, entre 2012 e 2018, e proposto um modelo para previsão de valores vencedores em licitações [28]. A escassez de ferramentas à disposição do governo e o efeito da competitividade entre empresas no preço pago motivou o uso do algoritmo de floresta aleatória para criação de um modelo de regressão utilizando 15 características preditoras. O objetivo do modelo era previsão de preços para itens homologados e, assim, auxiliar na tomada de decisão durante o processo licitatório.

Por fim, dando continuidade ao trabalho desenvolvido pelos autores desta pesquisa ao longo do mestrado [37], a pesquisa desenvolvida por De Lima *et al.* [24] avaliou a justiça e métodos de mitigação de injustiça em modelos de estimação de risco semelhantes aos utilizados por órgãos de controle. Essa avaliação de possíveis vieses considerou as esferas federal e municipal, bem como distinguiu o risco de empresas e contratos. Regressão logística, floresta aleatória, redes neurais, SVM e KNN foram os algoritmos utilizados na modelagem de risco.

3.1.3 Lances e Participantes em Leilão

O pregão eletrônico ^[11] opera como um leilão inverso. Isto é, quem propõe o menor lance e está de acordo com as demandas do proponente é contemplado. Esta é uma das diversas modalidades que a AP dispõe para contratar bens ou serviços privados. A detecção de figurantes (participantes cujo objetivo é influenciar o resultado do pregão, prejudicando o comprador e/ou vendedor) ou de lances possivelmente fraudulentos é tema abordado na literatura e consiste na estimação da chance de que um participante do leilão ou um lance sejam irregulares.

Na literatura referente a leilões ordinários, Rebouças *et. al.* ^[45] identificaram indicadores que pudessem ser portados ao contexto dos pregões eletrônicos. Utilizando dados do portal de compras do Governo Federal – o antigo ComprasNet ^[12] (atual Compras.gov) –, os autores propuseram um conjunto de indicadores relacionados a situações suspeitas para fundamentar a classificação de um lance como possivelmente irregular, partindo de um figurante. Após atribuir pesos a esses indicadores, foi gerado um conjunto de dados contendo figurantes e lances artificiais. Assim, a partir do cálculo do score dos lances, utilizando as características propostas, foi possível mensurar o risco dos lances e verificar sua autenticidade.

Ainda que utilizando dados da indústria (contexto privado), as contribuições trazidas por Anowar *et al.* ^[2] podem ser aplicadas à AP. Esse estudo avaliou a utilização de técnicas para balanceamento entre classes na identificação de lances possivelmente fraudulentos em leilões do eBay ^[13]. A partir de dados sobre o leilão de um videogame, foi empregado agrupamento hierárquico para rotulação desses lances em grupos significativos. O número de observações em cada classe de risco foi balanceado por meio de cinco técnicas distintas: SMOTE, SmoteENN, Smote-TomekLink, NearMiss, ClusterCentroids. Após balanceamento, o conjunto de dados foi utilizado para modelagem utilizando o algoritmo SVM. O modelo com balanceamento via SmoteENN demonstrou melhores resultados a partir da análise do F_1 .

¹¹ Pregão Eletrônico - Portal de Compras do Governo Federal

¹² ComprasNet - Governo Federal

¹³ eBay

3.1.4 Servidores Públicos

Desenvolvido na Diretoria de pesquisa e informações estratégicas da CGU e com o objetivo de prever o risco de corrupção de servidores públicos federais, Carvalho *et al.* [15] utilizaram dados referentes a filiação partidária de servidores públicos. Esse trabalho oportuniza uma comparação entre algoritmos de classificação selecionados e um modelo construído na prática por especialistas da área. Supondo uma ligação moral entre as esferas políticas e públicas, os autores deste artigo fazem uso de modelos com redes bayesianas, SVM, florestas aleatórias e redes neurais artificiais tradicionais com retroalimentação para classificação das observações. Para rotulagem dos dados de treinamento, foram usados registros disponíveis no sistema de administração de pessoal do governo e filtrados os casos em que servidores foram expulsos devido à corrupção. O modelo criado com floresta aleatória demonstrou os melhores resultados — ainda que sua precisão tenha sido semelhante ao modelo utilizado pelos especialistas (86%), a revocação obteve um valor aproximadamente duas vezes mais alto (32%).

3.1.5 Unidades Administrativas

No intuito de otimizar estratégias de controle, a CGU utilizou modelos probabilísticos (*Naïve Bayes*, *Tree Augmented Naïve Bayes*, e *Attribute Weighted Naïve Bayes*) para estimar o risco de corrupção em unidades administrativas federais [14]. A partir de dados sobre as unidades administrativas e sobre servidores expulsos em decorrência de comportamento corrupto, foi possível criar uma base de treinamento composta por unidades corruptas ou não corruptas. O modelo *Naïve Bayes* obteve melhor desempenho, atingindo uma AUC-ROC de 0,76.

Já no Tribunal de Contas da União (TCU), foi experimentada uma estratégia para estimar o risco combinado de corrupção e conluio entre unidades governamentais e empresas privadas na execução de contratos [3]. O algoritmo *Naïve Bayes* foi utilizado numa abordagem não-supervisionada e sua validação foi realizada através do conhecimento e opinião de especialistas na área.

3.2 Identificação de Contexto

Na publicização de dados em portais de transparência ou sistemas de gerenciamento de recursos, via de regra, os gastos realizados pela AP são categorizados por propósito ou contexto. Em alguns estados, quando disponível, uma estratégia possível para identificação do tipo/contexto do gasto é acessar despesas semelhantes através da sua função e/ou subfunção orçamentárias — que indicam de qual fundo de recursos o valor gasto é proveniente. Essa identificação de contexto tem potencial de favorecer a transparência e o controle social possibilitando buscas proativas por temas de competência ou, ainda, comparações em operações semelhantes. Entretanto, um entrave no uso da estratégia relacional de função ou subfunção orçamentária do gasto é a frequente indisponibilidade, incorretude ou incompletude da informação.

O Tribunal de Contas do Estado de Goiás utilizou os algoritmos SVM, naive bayes e regressão logística para categorização de gastos do Sistema de Informações Orçamentárias e Financeiras Integradas (SIOFI) de acordo com seu propósito [23]. Através de uma modelagem supervisionada do texto descritivo do gasto (possibilitada pela rotulagem manual de parte dos dados de entrada), foi possível identificar 124 despesas (3,1 milhões de dólares) relacionadas à propaganda/publicidade não divulgadas previamente ao público.

O uso do texto descritivo de itens também foi útil na geração de uma base de preços de referência, ferramenta de necessidade básica em unidades gestoras para realizar cotações e comparações de preço diariamente [13]. Na nota fiscal, são detalhados, em texto semi-estruturado, os códigos identificadores dos itens descritos, o que contribui para a categorização. Através desse dado, os autores desenvolveram uma metodologia para agrupamento de gastos em itens semelhantes utilizando KNN.

3.3 Considerações

É relevante pontuar que o acesso a artigos e trabalhos iniciais sobre determinados temas só foi possível graças ao contato com servidores públicos que, gentilmente, concederam acesso às suas produções que, ocasionalmente, estão disponíveis apenas ao público interno dos órgãos. Ainda, diversos trabalhos contam com a participação direta de membros de órgãos

governamentais. Este trabalho de tese contou com a generosa colaboração de servidores do Ministério Público da Paraíba e Tribunal de Contas do Estado da Paraíba, entretanto foi desenvolvido e orientado por membros da sociedade civil.

Os trabalhos citados nessa revisão de literatura representam grande contribuição para o estado da arte tanto na área computacional, como no controle social. No primeiro, inovando na aplicação de técnicas em diversos contextos, propondo novas abordagens e produzindo conhecimento acerca dos comportamentos estudados. No último, auxiliando as entidades fiscalizadoras a migrarem para uma postura proativa, de modo a antever casos de corrupção, de fraude e otimizando suas próprias metodologias e uso de recursos.

Em se tratando de trabalhos que fazem uso AM para revisão de algum comportamento de risco, é pertinente elencar decisões conceituais cujo objetivo seja estimar o **risco de contratos** públicos. Como visto, há casos em que: (i) são utilizados rótulos de classes referentes a comportamentos prévios de empresas; (ii) são utilizadas observações acerca de empresas; ou (iii) são utilizados rótulos referentes a meras indicações de risco, e não de comportamentos comprovadamente irregulares no passado.

No primeiro e segundo casos, no entendimento da autora deste trabalho, é fundamental que haja a distinção entre as entidades de contrato e empresa, uma vez que a multiplicidade do par empresa-contrato é de 1:N. Isto é, uma mesma empresa pode estar associada a um ou mais contratos públicos. Sendo assim, para estimar o risco de um contrato, aponta-se para o menor grão do processo de contratação, ao passo que, ao estimar o risco de uma empresa, será necessária uma varredura em todos os contratos associados à mesma. No terceiro caso, entende-se que o uso das indicações de risco é válido e fundamental, quando utilizado como uma característica preditora (variável independente), uma vez que um comportamento passado não pode ser constatado em suposição, senão em um fato concreto.

No panorama brasileiro, comportamentos de risco são predominantemente abordados considerando entidades da esfera federal. Com exceção da pesquisa desenvolvida por De Lima [24], cujo foco foi análise de justiça e viés em modelos previamente propostos, não foram encontrados outros estudos abordando risco em nível municipal ou estadual. Além disso, não há evidências de uma avaliação experimental da prática na estimação do risco em contratos e empresas. Por fim, especificamente para os municípios do estado da Paraíba, não há registro do uso de AM para estimação de risco em contratos públicos. Um sumário dos

estudos apresentados está disponível no Apêndice A¹⁴.

O uso de características latentes na modelagem do risco de rescisão contratual, bem como de técnicas para explicar as previsões do modelo ou compará-las com a prática do controle, não foi identificado na literatura, seja no contexto federal, estadual ou municipal. O preenchimento dessas lacunas representa uma contribuição valiosa, especialmente no nível de contratos municipais, onde frequentemente não há dados estruturados ou rotulados para análise qualificada de dados ou treinamento de modelos de AM.

O objetivo deste trabalho é estender o uso de AM na esfera pública ao contemplar contratos municipais do estado da Paraíba. Como primeiro passo, é avaliado um método prático para estimação de risco de empresas nesse estado, explorando suas deficiências e utilidades. Essa análise fornece argumentos experimentais para, em seguida, enquadrar a estimação de risco no nível de contrato. A partir disso, são experimentadas técnicas de balanceamento de classes, novos conjuntos de características e estratégias para modelagem do risco contratual. Em seguida, as contribuições das características relevantes para a composição do risco de rescisão são analisadas de modo geral (para todo o conjunto de contratos) e, também, em uma amostra dos dados (para os contratos corretamente classificados pelo modelo). Por fim, compara-se a semântica das características mais relevantes na previsão do risco dos contratos pelo modelo com pareceres técnicos de auditores, visando validar o constructo de risco de rescisão contratual proposto neste trabalho.

¹⁴ A Tabela A.1 lista os dados e as técnicas utilizadas, enquanto a Tabela A.2 detalha os métodos e as conclusões.

Capítulo 4

Materiais e Métodos

Este capítulo descreve os métodos e materiais utilizados nos experimentos e análises executados nos capítulos seguintes deste documento. Oportunamente, esta pesquisa combina a avaliação experimental de modelos, a aplicação de técnicas do estado da arte em AM e a caracterização do comportamento de risco de rescisão contratual no contexto municipal.

4.1 Fontes de Dados

De modo geral, foram aproveitados dados sobre contratações públicas e execução da despesa pública, registros empresariais, sanções empresariais anteriores, resultados de auditorias, entre outros. A partir de filtros, transformações e cruzamentos entre as bases de dados abaixo listadas, foram construídos os conjuntos de dados de contratos utilizados como base na modelagem do risco de rescisão e análises descritas nos próximos capítulos.

SAGRES

Como descrito na Subseção [2.1.3](#), o banco de dados do sistema SAGRES contém dados sobre o processo de contratação, sobre a relação histórica das empresas com a AP, informações cadastrais de servidores e fornecedores, empenhos, pagamentos e outros documentos resultantes do processo de contratação e despesa pública estadual e municipal da Paraíba. A este trabalho, foi útil o módulo SAGRES Municipal, dividido em 294 tabelas, do qual foram extraídos dados sobre propostas, licitações, contratos, empenhos, fornecedores e pagamentos.

Receita Federal do Brasil

A Secretaria Especial da Receita Federal (RF) do Brasil é responsável por manter um cadastro de todas as pessoas jurídicas (empresas) em operação em todo o Brasil, o Cadastro Nacional da Pessoa Jurídica (CNPJ). Nesse cadastro estão contidas as datas de abertura, o porte, a situação de operação da empresa, entre outros dados. Neste estudo filtramos apenas dados sobre empresas em contratação com a AP paraibana. Esses dados tornaram-se, recentemente, disponíveis ao público para descarga massiva e consulta por meio do portal de dados abertos do Governo Federal¹⁵.

Matriz de Risco de Empresas

Essencialmente, em qualquer contexto, uma matriz de risco condensa um conjunto de regras baseadas em comportamentos de risco, de maneira a detectar ou antever possíveis irregularidades. Originalmente elaborada por servidores públicos da Controladoria Geral de Pernambuco e adotada pelo Tribunal de Contas da União, além de outros órgãos, a matriz de risco (MR) utilizada neste trabalho reúne regras baseadas no comportamento licitatório e dados cadastrais de empresas em relacionamento com a AP.

A MR foi elaborada a partir da experiência prática de servidores públicos, com a intenção materializar tipologias de risco. Essa é uma dentre as ferramentas utilizadas pelo Ministério Público da Paraíba (MPPB) e outros órgãos, para estimação de risco de empresas e, até onde foi possível verificar, não foi previamente avaliada experimentalmente.

A instância utilizada da ferramenta contém atributos estratégicos sobre mais de 29.000 empresas em contrato com os órgãos públicos paraibanos. A matriz computa 57 variáveis que descrevem o comportamento passado das empresas, dados de registro, possíveis sanções prévias, composição societária e informações sobre os sócios, entre outros dados. O sumário (score de risco) dessas variáveis é representado por uma soma ponderada desses 57 valores, que têm seus pesos definidos manualmente por especialistas.

¹⁵Cadastro Nacional da Pessoa Jurídica – Governo Federal.

Tramita

O Tramita¹⁶, Sistema de Tramitação de Processos e Documentos do Tribunal de Contas do Estado da Paraíba, permite o acompanhamento da situação das contas e tramitação dos processos dos municípios. Nesse sistema ficam armazenados as mutações de estado dos contratos da AP paraibana. As mutações podem caracterizar-se nos seguintes tipos: aditivo, apostilamento, retomada, ordem de serviço, paralisação, rescisão, suspensão, impedimento ou sustação.

Dentre essas, foram selecionadas as mutações relativas aos comportamentos que pudessem, de alguma maneira, encerrar ou interromper o curso esperado da execução do contrato para que fosse possível operacionalizar o conceito e constructo de risco. Assim, a partir de agora referidas por apenas "rescisões", foram utilizadas mutações contratuais referentes a paralisações, rescisões, suspensões, impedimentos e sustações. Na Figura 4.1, observamos as ocorrências dessas mutações de interesse entre os anos de 2014 e 2019.

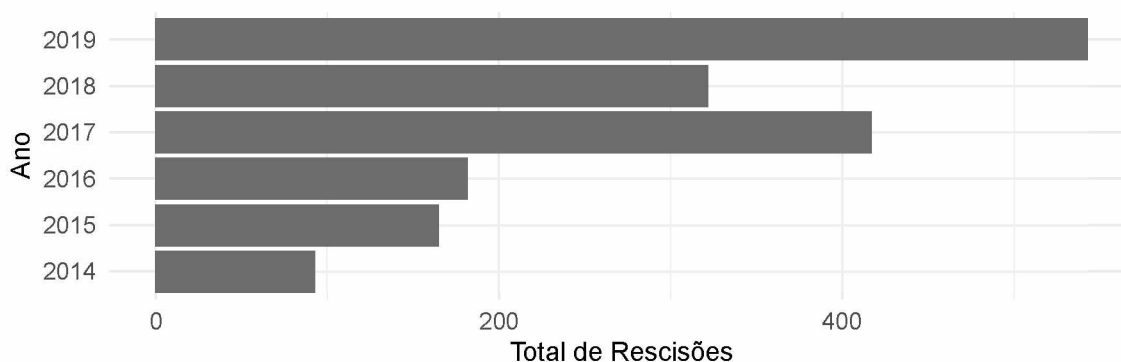


Figura 4.1: Total de mutações contratuais de interesse ao longo dos anos. Dados: Tramita. Fonte: elaboração própria.

A partir da figura, vê-se que o número de mutações das categorias de interesse é crescente ao longo dos anos. Contudo, não é possível afirmar se esse aumento se dá em decorrência do aumento do próprio número de contratos firmados a cada ano, ou devido ao exercício mais intenso das atividades de controle.

¹⁶Tramita TCE-PB.

4.2 Operacionalização do Risco de Rescisão

Medir o risco associado a algo pode ser definido como estimar o grau de incerteza de que um determinado evento ocorra ou não. Por exemplo, quando alguém solicita um empréstimo a uma instituição bancária, esta realiza uma avaliação a fim de verificar se o solicitante receberá ou não o valor solicitado. Pode-se dizer que, nessa análise de adimplência/inadimplência, é avaliado o crédito do solicitante ou o risco que o banco corre ao garantir o empréstimo do dinheiro. Nesse exemplo simplório, é intuitivo entender o porquê de características como histórico de pagamentos, dados sobre bens, receitas e despesas serem úteis à decisão de concessão do empréstimo.

A existência do risco, tal como sua intensidade, está condicionada a fatores específicos do domínio em questão. Assim, para analisar o risco em outros cenários análogos ao exemplo anterior, é necessário observar e investigar duas dimensões fundamentais: o evento que se deseja observar e as características do cenário em questão que podem influenciar no crescimento ou diminuição da chance de ocorrência desse evento. A estratégia de estimar a chance de ocorrência de um evento com base em características, eventos anteriores e seus resultados tem sido utilizada por diversos setores de maneira a lidar com grandes volumes de solicitações de análise [2; 30]. Nos Capítulos 5 e 6 são abordadas a avaliação experimental do risco associado a empresas e a contratos públicos. Em cada capítulo são definidas as características a partir das quais é modelado o risco.

A literatura disponível acerca da previsão de ocorrência de corrupção, interrupções, ou quaisquer outros eventos indesejados durante a execução do contrato, tipicamente, recorre a duas estratégias para operacionalização desses comportamentos para modelagem. Na primeira estratégia, mais utilizada, faz-se uso da ocorrência passada de eventos que representam irregularidades ou comportamentos sabidamente reprováveis para viabilizar a rotulagem do conjunto de dados. Na segunda estratégia, faz-se uso de indicações de risco – que não são necessariamente ações irregulares –, para rotulagem do conjunto de dados. Em ambos os casos, considerando uma tarefa de classificação supervisionada do risco.

A pressuposição da existência de registros históricos suficientemente consistentes e acessíveis, tendo em vista, em muitas cidades, a pouca transparência dos gastos públicos ou limitações de governança institucional, são implicações relevantes à primeira estratégia. Ade-

mais, a detecção parcial (subnotificação) dos casos sabidamente irregulares pode produzir um viés inerente à criação desse conjunto de rótulos, impactando negativamente a capacidade preditiva de qualquer modelo de risco. Apesar dessas limitações, essa abordagem de operacionalização do risco considera a real ocorrência de eventos, por lei, sabidamente irregulares ou prejudiciais ao processo de contratação. Por outro lado, o uso de indicações de risco para a operacionalização do conceito de risco está principalmente limitado à consideração subjetiva do que constitui um comportamento arriscado e que merece uma inspeção mais detalhada, uma vez que esse comportamento é uma mera suposição.

Abaixo, detalha-se como é posto em prática o conceito de risco para empresas e contratos, uma vez que diferentes órgãos da AP, praticantes e pesquisadores têm diferentes formas de tornar palpável a abstração do risco para entidades. O processo de avaliação do risco de que uma empresa seja fantasma ou que um contrato tenha um alto risco de corrupção ou fraude, por exemplo, tomará rumo de acordo com o conhecimento prévio e experiência de cada praticante.

4.2.1 Risco em Empresas

Como visto anteriormente, órgãos de controle fazem uso de estratégias de detecção de risco também a nível das empresas contratadas. As tipologias de risco da MR auxiliam na priorização de casos mais críticos através de scores de risco. Essas tipologias podem ser comparadas, em analogia, às características sobre o solicitante do empréstimo exemplificado acima, e caracterizam a ocorrência de um evento de interesse comum: o risco de irregularidade. Nesse caso, a sumarização destas tipologias torna tangível o conceito de risco para cada empresa listada.

Para a avaliação da MR no Capítulo 5, vê-se seu comportamento desde o ponto de vista da avaliação de um modelo de AM: o ranking gerado pela MR servirá como a "previsão" do modelo/matriz, e a informação de rescisão vinda do Tramita atuará como o rótulo real dessa observação. Dessa maneira, o conceito de risco de uma empresa é associado à existência de contratos previamente rescindidos associados a ela. Assim, *uma empresa de alto risco é toda e qualquer empresa que tenha, relacionada a si, pelo menos um contrato rescindido.*

4.2.2 Risco em Contratos

Similar ao risco de empresas descrito acima, o conceito de risco para contratos públicos é materializado através de construção de um conjunto de variáveis que caracterizam o próprio contrato em situações que o relacione ao evento de interesse que é utilizado como rótulo – nesse caso, que possa relacionar o contrato ao comportamento futuro de rescisão contratual. Dessa maneira, *um contrato de alto risco é todo e qualquer contrato que tenha incorrido em uma das mutações de interesse.*

É importante destacar que, em ambos os conjuntos de dados mencionados acima, englobando empresas e contratos, a classe de baixo risco que rotula parte das observações não denota a ausência de qualquer irregularidade, mas sim que essas não foram previamente inspecionadas para atribuição de um rótulo.

Ainda, outras fontes de dados podem ser utilizadas como formas de operacionalizar a abstração do risco, através de rotulagem das entidades a serem classificadas. O Cadastro Nacional de Empresas Inidôneas ou Suspensas (CEIS), o Cadastro Nacional de Empresas Punidas (CNEP), o Cadastro Nacional de Condenados por Ato de Improbidade Administrativa e por Ato que Implique Inelegibilidade (CNCIAI), o Cadastro de Expulsões da Administração Federal (CEAF) e o Cadastro de Fornecedores Impedidos de Licitar e Contratar com a Administração Pública Estadual da PB (CAFIL) são exemplos de outras fontes de dados de comportamentos sancionáveis. Entretanto, a depender do contexto do estudo, é importante verificar o grau de subnotificação desses cadastros, a fim de maximizar a performance do modelo e atingir o máximo possível de balanceamento entre as classes de risco. O CEIS, por exemplo, é a fonte de dados mais comum em esforços anteriores na literatura disponível para a esfera federal de governo. Entretanto, é uma fonte de dados severamente subnotificada no contexto municipal.

4.3 Avaliação em Cenários Desbalanceados

Modelos de classificação treinados a partir de conjuntos de dados desbalanceados requerem métricas de avaliação adequadas para evitar interpretações equivocadas sobre sua performance. Nesse cenário, estudos anteriores adotam a AUC-ROC, uma métrica amplamente utilizada em AM, cujo objetivo é ser maximizada para avaliar a capacidade do modelo em

distinguir entre classes positivas e negativas.

A AUC-ROC é definida como a área sob a curva que representa a relação entre a Taxa de Verdadeiros Positivos (*True Positive Rate* – TPR), também conhecida como Revocação (*Recall*), e a Taxa de Falsos Positivos (*False Positive Rate* – FPR), para a avaliação de classificadores considerando um intervalo como limiares da classificação [7]. Respectivamente, TPR e FPR são definidas pelas seguintes fórmulas:

$$TPR = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (4.1)$$

$$FPR = \frac{\text{False Positives}}{\text{False Positives} + \text{True Negatives}} \quad (4.2)$$

Entretanto, em um cenário de severo desbalanceamento, o número de Verdadeiros Negativos (*True Negatives*) é tão grande que torna a FPR irrelevante, encobrindo possíveis ganhos na FPR, na avaliação do modelo. Como alternativa, é preferível que a métrica penalize modelos que prevejam todas, ou grande parte, das observações como classe majoritária, enquanto também seja sensível a Falsos Positivos. Em nosso contexto, especificamente, uma métrica apropriada consideraria Falsos Positivos tão importantes quanto Verdadeiros Positivos e Falsos Negativos. O aumento de Falsos Positivos demanda grandes despesas em capital humano e financeiro para sustentar investigações, assim como o aumento de Falsos Negativos enfraquece o propósito de identificar entidades de alto risco.

Após investigação e experimentação com diversas métricas, o estado da arte [7; 22] aponta que o Coeficiente de Correlação de Matthews (*Matthews Correlation Coefficient* – MCC) e a Área Sob a Curva de Precisão-Revocação (AUC-PR) são os indicadores de performance mais apropriados para cenários de severo desbalanço entre classes, devido ao comportamento não enviesado em relação ao grupo majoritariamente representado. Entretanto, sendo a primeira métrica pouco utilizada na literatura de estimação de risco, será utilizada a AUC-PR para avaliação e comparação dos modelos propostos.

A AUC-PR é uma métrica difundida na literatura e capaz de discernir o desempenho de modelos de risco em cenários desequilibrados. Essa métrica é definida como a área formada pela relação entre os valores de Precisão (Equação 4.3) e Revocação (TPR), também considerando a avaliação do classificador em diferentes limiares de classificação. Quanto maior o valor para ambas as medidas, maior será a AUC-ROC – o aumento de Verdadeiros Positivos diminui Falsos Negativos e Falsos Positivos.

$$\text{Precisão} = \frac{\text{Verdadeiros Positivos}}{\text{Verdadeiros Positivos} + \text{Falsos Positivos}} \quad (4.3)$$

Uma importante distinção entre AUC-ROC e AUC-PR está no seu limiar base de classificação (que é frequente e comumente utilizado como comparativo de performance). Enquanto, para a AUC-ROC, um classificador aleatório tem performance de 0,5 em um conjunto de dados balanceado; para a AUC-PR, um classificador aleatório teria uma performance igual ao percentual de casos da classe positiva no conjunto de dados, como definido em 4.4.

$$\text{Linha base AUC-PR} = \frac{\text{Num. Obs. Classe Pos.}}{(\text{Num. Obs. Classe Pos.} + \text{Num. Obs. Classe Neg.})} \quad (4.4)$$

Para avaliação dos modelos apresentados neste trabalho, é computado o intervalo com 95% Confiança (IC), com 4.000 reamostragens *bootstrap* da amostra utilizada como conjunto de dados de testes para reportar e interpretar resultados. Assim, serão realizadas inferências a partir do intervalo construído e sua estimativa [19].

4.4 Normalização min-max

A normalização min-max é uma transformação que permite escalar um conjunto de valores para um determinado intervalo, relacionando o valor mínimo do conjunto ao limite inferior do intervalo e o valor máximo ao limite superior do intervalo.

$$X' = \frac{(X - X_{min})}{(X_{max} - X_{min})} \quad (4.5)$$

A fórmula 4.5 define a operação. X é o valor corrente a ser transformado, X_{min} e X_{max} são os valores mínimo e máximo do conjunto, respectivamente. Por fim, X' é o valor normalizado resultante. Nesse estudo, a normalização min-max será utilizada para atribuição de pesos de relevância com base no valor do contrato durante o processo de aprendizagem do modelo.

4.5 Infraestrutura

Para pré-processamento de dados foi utilizada a linguagem e ambiente de computação R (versão 3.5.1), bibliotecas: RPostgres¹⁷, DBI¹⁸, Config¹⁹, RMySQL²⁰, Tidyverse²¹, Here²² e Tibble²³. Para treinamento e avaliação dos modelos, foi utilizada a linguagem Python (versão 3.10.4) e bibliotecas: pandas²⁴, numpy²⁵, scikit-learn²⁶, xgboost²⁷, matplotlib²⁸ e pickle²⁹. Bem como também R, com as bibliotecas: Tidymodels, Caret³⁰, Unbalanced³¹, Mccr³², GridExtra³³, Ggbeeswarm³⁴ e ggplot2³⁵. Para o cálculo e a geração dos intervalos

¹⁷<https://www.rdocumentation.org/packages/RPostgres/versions/1.1.1>

¹⁸<https://www.rdocumentation.org/packages/DBI/versions/0.5-1>

¹⁹<https://www.rdocumentation.org/packages/config/versions/0.3>

²⁰<https://www.rdocumentation.org/packages/RMySQL/versions/0.10.17>

²¹<https://www.rdocumentation.org/packages/tidyverse/versions/1.2.1>

²²<https://www.rdocumentation.org/packages/here/versions/0.1>

²³<https://www.rdocumentation.org/packages/tibble/versions/1.4.2>

²⁴<https://pandas.pydata.org/pandas-docs/version/2.2.0rc0/whatsnew/v2.2.0.html>

²⁵<https://numpy.org/doc/2.1/release/1.26.4-notes.html>

²⁶https://scikit-learn.org/stable/whats_new/v1.4.html

²⁷https://xgboost.readthedocs.io/en/release_2.0.0/install.html

²⁸<https://matplotlib.org/3.8.2/>

²⁹<https://docs.python.org/3/library/pickle.html>

³⁰<https://www.rdocumentation.org/packages/caret/versions/6.0-84>

³¹<https://www.rdocumentation.org/packages/unbalanced/versions/2.0>

³²<https://www.rdocumentation.org/packages/mccr/versions/0.4.4>

³³<https://www.rdocumentation.org/packages/gridExtra/versions/2.3>

³⁴<https://www.rdocumentation.org/packages/ggbeeswarm/versions/0.5.3>

³⁵<https://www.rdocumentation.org/packages/ggplot2/versions/3.2.0>

de confiança, foi utilizada a linguagem R, com as bibliotecas: Broom³⁶ e Boot³⁷.

³⁶<https://www.rdocumentation.org/packages/broom/versions/0.5.2>

³⁷<https://www.rdocumentation.org/packages/boot/versions/1.3-22>

Capítulo 5

Estimação de risco de empresas: prática e eficácia

Apesar da experimentação e eventuais aplicações de técnicas de IA por órgãos públicos em desafios cotidianos, como visto no Capítulo 3, os sistemas para auxílio em investigações, fiscalizações ou auditorias frequentemente utilizam soluções computacionais mais simples e que não envolvem a identificação e aprendizado automático de padrões. Isso pode ocorrer devido à complexidade de rotulagem dos dados, ao custo de implementação e manutenção, ou mesmo devido à baixa adesão ou resistência das áreas interessadas ao uso das ferramentas.

O risco de ocorrência de um determinado evento pode ser caracterizado, operacionalizado e mensurado através de padrões de ocorrência identificados na experiência de especialistas em uma área ou negócio. No contexto da contratação de empresas pela AP, empresas que possuem sócios beneficiários de programas de assistência social ou que firmam quantidades de contratos em desacordo com seu porte estrutural, financeiro ou técnico podem levantar suspeitas de irregularidades. Nesse sentido, essas regras baseadas em comportamentos de risco são chamadas de trilhas ou tipologias de risco. De maneira análoga, outros aspectos podem ser analisados para criar padrões de comportamento incomuns, os quais podem ser relevantes à participação dessa empresa em processos licitatórios ou ao cumprimento de suas obrigações contratuais de acordo com a lei e o contrato.

5.1 Matriz de risco

A MR é composta por variáveis numéricas que materializam características consideradas passíveis de atenção para cada empresa em contrato com a AP. Essas variáveis englobam informações sobre histórico de prestação de serviços, composição societária e dados cadastrais das empresas, quadro de funcionários, porte/capital social, localização, entre outros atributos. Assim, a partir da atribuição de pesos para as características históricas e cadastrais, é construído um score único de risco para cada empresa.

Gerado em 17 de junho de 2017, o arquivo da MR é composto por pouco mais de 29.000 registros, com 62 variáveis, dentre as quais 57 são tipologias de risco. Um ranking do risco associado às empresas pode ser gerado a partir da ordenação inversa do score de risco – quanto maior o score, maior o risco de irregularidade.

Neste capítulo, é quantificada a eficácia prática da MR com o objetivo de avaliar experimentalmente a própria ferramenta e fornecer uma linha de base para comparação com os modelos de AM posteriores. Além disso, também busca-se orientar desenvolvedores e usuários sobre o impacto da MR com base no método avaliativo desenvolvido. Cabe ressaltar que a MR avaliada representa uma instância específica, um retrato momentâneo da massa de dados disponível, sendo apenas uma das várias ferramentas utilizadas para atividades de controle.

5.2 Filtragem e rotulagem de dados

Ao considerar a inviabilidade prática de um órgão de controle avaliar todas as 29 mil empresas da MR, o escopo de avaliação da ferramenta foi reduzido a uma amostra de 15 mil empresas, selecionadas com base nos maiores scores de risco, de modo a tornar a análise mais próxima da realidade.

Para avaliar o modelo treinado em uma tarefa de classificação, o rótulo previsto pelo modelo é comparado ao rótulo previamente estabelecido para cada observação. Nesse cenário, as rescisões contratuais do Tramita foram utilizadas para rotular cada observação da MR, enquanto o score de risco produzido pelas tipologias desempenhou o papel de previsão do classificador.

Desse modo, duas classes de risco foram criadas com base nas rescisões: empresas com pelo menos um contrato rescindido foram rotuladas como de alto risco, enquanto empresas sem contratos rescindidos foram rotuladas como de baixo risco. O conjunto de dados final contém 14.595 empresas de baixo risco e 405 empresas de alto risco.

A "classificação" da MR foi gerada a partir de subconjuntos de dados nas Top N posições no ranking de score de risco, com o tamanho dos subconjuntos iguais a 10, 50, 100 e 500 empresas. Assim, é considerado que todas as empresas em cada um dos subconjuntos possuem algum indício de irregularidade (alto risco), enquanto as demais não possuem indício de irregularidade (baixo risco).

Na Figura 5.1 é possível observar a distribuição do score (Total Tipologia) em relação às classes de alto e baixo risco, com destaque para a diferença de tamanho entre as classes de risco.

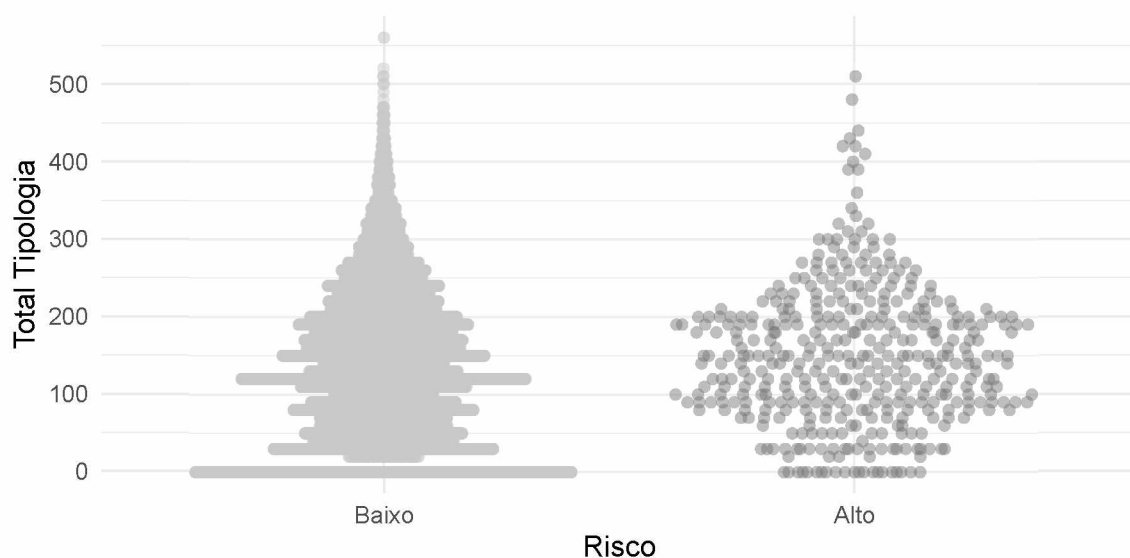


Figura 5.1: Distribuição do score de risco da MR por classe. Fonte: elaboração própria.

5.3 Avaliação da matriz de risco

A classe de empresas consideradas de alto risco representa 2,7% (405) da amostra de 15 mil observações. Considerando a AUC-PR como métrica de avaliação apropriada em um cenário desbalanceado, um classificador aleatório para esse conjunto de dados teria uma performance

de 0,027 (como detalhado na Seção 4.3). Esse valor servirá de referência para avaliação da classificação do risco nos subconjuntos de empresas.

Tabela 5.1: Avaliação da Matriz de Risco

Conj.	VP	VN	FP	FN	Precisão	Revoc.	AUC-ROC	AUC-PR IC
10	2	14,587	8	403	0.200	0.004	0.502	[-0.035, 0.218]
50	4	14,549	46	401	0.080	0.009	0.503	[0.008, 0.093]
100	8	14,503	92	397	0.080	0.019	0.507	[0.028, 0.093]
500	13	14,108	487	392	0.026	0.032	0.499	[0.026, 0.055]

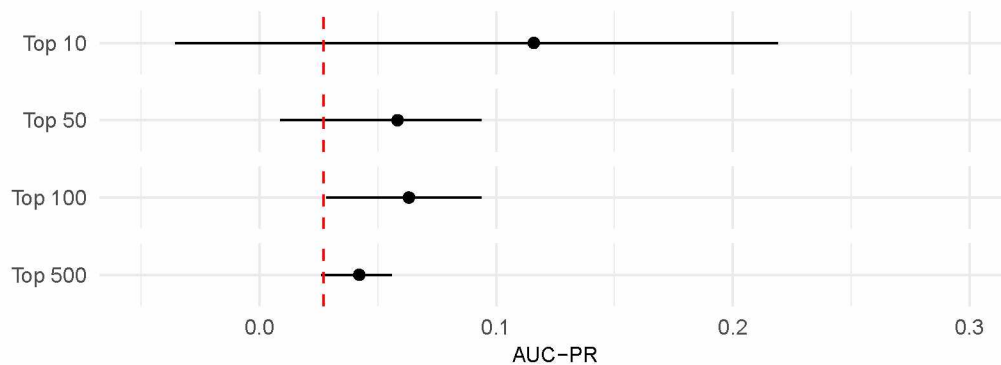


Figura 5.2: Intervalos de Confiança (AUC-PR) - Matriz de Risco

A Tabela 5.1 exibe as métricas de avaliação para os cinco subconjuntos de empresas com maior pontuação de risco. Quanto mais próximo o AUC-PR de um subconjunto de empresas esteja do AUC-PR do classificador aleatório, menor será a capacidade de distinção entre as classes de risco nesse recorte e, conseqüentemente, menor será a credibilidade do score de risco na distinção entre possíveis casos irregulares. Na Figura 5.2, pode-se comparar o desempenho dos subconjuntos da MR com o desempenho de um classificador aleatório hipotético (representado pela linha vermelha vertical).

A partir dos ICs, nota-se que, à medida que o número de empresas considerado cresce, a distribuição dos valores de AUC-PR restringe-se a uma menor amplitude — o que representa um intervalo mais acurado. Contudo, apesar de uma menor variabilidade entre os valores de AUC-PR possíveis, o desempenho na seleção de empresas com potencial de alto risco tende

a decair (piora na classificação).

5.4 Discussão

Para os subconjuntos de 10 e 50 empresas, com 95% de confiança, não há evidências conclusivas de que o desempenho da MR seja superior a um palpite aleatório, podendo ser pior. Porém, é plausível que a performance de classificação desses subconjuntos tenha vantagem em relação ao classificador aleatório. A partir das 100 e 500 empresas, o desempenho desses subconjuntos pode ser equivalente ou superior ao classificador aleatório. No melhor cenário (limite superior do intervalo), esses subconjuntos alcançariam desempenhos pelo menos três (0,093) e duas (0,055) vezes superiores ao classificador aleatório, respectivamente.

A precisão informa sobre o quão correta é a seleção de empresas positivamente suspeitas feita pela classificação. Já a revocação fornece a informação de quantas empresas de alto risco a MR está classificando corretamente em relação ao total de empresas de alto risco. À medida que o tamanho do subconjunto de empresas cresce, a precisão diminui e a revocação aumenta. Sob esse viés, a MR pode alcançar desempenhos aceitáveis quando são utilizados subconjuntos de empresas suficientemente abrangentes, o que conflita com o objetivo intrínseco da priorização de risco: minimizar o escopo de trabalho para reduzir os recursos despendidos.

Considerando o propósito dos scores de risco da MR, quanto menor fosse o subconjunto de empresas (que estariam associadas aos maiores scores de risco), maiores seriam também a precisão e a revocação do subconjunto. Entretanto, apesar da previsão de risco acertada para uma pequena parcela das empresas – o que pode sinalizar certa convergência entre as tipologias de risco e o descritor de risco de rescisão contratual –, o número de empresas falsamente classificadas como alto risco representa uma fraqueza da ferramenta. Por exemplo, em um cenário de análise dos subconjuntos de 100 e 500 empresas e uma situação de recursos financeiros e humanos limitados, ainda seriam inspecionadas 92 e 487 empresas (e, talvez, seus contratos relacionados) desnecessariamente.

A relevância de métodos de estimação de risco está justamente no direcionamento de esforços de fiscalização (humanos e financeiros) de maneira mais eficaz, idealmente para casos em que já haja sinais de comportamentos irregulares. Nesse caso, uma postura proativa

acerca do comportamento contratual das empresas poderia assegurar o investimento público; entretanto, como visto, a maior parcela das empresas ainda seria examinada em vão. Em um contexto realista, seria mais comum que possíveis irregularidades ocorressem em níveis mais granulares como, por exemplo, nos contratos. Afinal, o histórico de relação de uma empresa com a AP é marcado por contratos celebrados com a finalidade de provimento de bens ou serviços.

Diante dessa necessidade, a modelagem e avaliação de estratégias de previsão de risco devem acontecer no nível do contrato ou entidade de grão similar ou menor que possa conduzir a intervenções práticas pelos órgãos responsáveis. No capítulo seguinte, na intenção de preencher essa lacuna, são propostos modelos de AM para estimar o risco de rescisão em contratos públicos.

5.5 Limitações

O objetivo dessa análise foi avaliar experimentalmente a eficácia da MR a partir do histórico de irregularidades nos contratos de suas respectivas empresas. Contudo, é relevante que limitações presentes nos conjuntos de dados utilizados ou no próprio método de sua construção sejam pontuadas, em especial, para fins de interpretação da análise e para aprimoramento dos esforços experimentais futuros.

Sub-representação do conjunto de rótulos

É importante ressaltar que o conjunto de rescisões é provável de subnotificação e sub-representação. A criação de um conjunto de rótulos completo envolveria a investigação de todos os contratos disponíveis, um cenário inviável, tendo em vista os recursos necessários para a tarefa. No conjunto de rótulos utilizado, estão presentes contratos que foram investigados previamente pelos órgãos competentes. Eles apresentaram algum indício de irregularidade, falha ou desvio em sua execução padrão e tiveram esses eventos documentados.

Desse modo, contratos não rescindidos não são necessariamente contratos idôneos, mas sim contratos que não passaram por escrutínio. Assim, empresas que não foram rotuladas como alto risco não são necessariamente isentas de risco. Elas são classificadas pela

MR como falsos positivos e podem exibir comportamentos atípicos que justificariam explorações mais detalhadas em seu comportamento, apesar de não terem contratos rescindidos associados a elas como rótulo.

Construção da MR

Como visto, as tipologias de risco foram desenvolvidas com base na experiência prévia de servidores públicos no exercício de suas funções. Naturalmente, a construção dessas tipologias reflete a opinião e o conhecimento tácito relacionado ao risco dessas pessoas, o que pode reproduzir vieses intrínsecos ao seu modo de trabalho e cosmovisão.

Ao usar a MR com ou sem algoritmos de AM, esses vieses podem impactar negativamente, de maneira desproporcional, empresas jovens, empresas de pequeno porte, produtores familiares que pleiteiem licitações por meio de programas sociais e outros grupos sensíveis de empresas cujas características encaixem-se em comportamentos tidos como arriscados [35]. Nas aplicações de AM, o uso de ferramentas como o Aequitas³⁸, por exemplo, pode auxiliar na identificação de grupos sensíveis e grupos de referência, além de auxiliar na mitigação de possíveis vieses na previsão dos modelos.

³⁸ Aequitas

Capítulo 6

Estimação do risco de rescisão em contratos

A avaliação apresentada no capítulo anterior permitiu mensurar o esforço de investigação necessário para identificar contratações de alto risco em determinadas empresas. Na prática, cada contrato firmado por essas empresas seria submetido a um escrutínio detalhado, visando identificar possíveis falhas ou irregularidades. Analisar um subconjunto de 10 empresas, das quais duas apresentam potencial risco, exigiria a avaliação de pelo menos 67 contratos associados a elas. Nesse contexto, a precisão de 0,2 na avaliação de risco de empresas cairia para 0,02 no nível de contratos, uma redução de dez vezes. Nos subconjuntos de 50 e 100 empresas, haveria 217 e 1.037 contratos a inspecionar, resultando em precisões de 0,01 e 0,007, respectivamente. Assim, estimar o risco de contratos com base no ranking de risco oferecido pela MR se mostra uma estratégia ineficaz. A ineficiência da estimação de risco para empresas e o uso promissor de AM para revisão de risco em entidades da contratação pública na literatura, detalhados no Capítulo 3, fundamentam a abordagem da previsão de risco de rescisão no contexto de contratos.

Neste capítulo, são apresentados os resultados da experimentação sobre a estimação de risco de rescisão em contratos públicos dos municípios da Paraíba. Além disso, avalia-se a eficácia de um algoritmo de aprendizagem do estado-da-arte para dados tabulares, compara-se estratégias comumente utilizadas para mitigação do desbalanço entre classes, e mensura-se o impacto da eventual indisponibilidade de conjuntos de dados na eficácia do modelo no contexto da contratação municipal.

6.1 Método

6.1.1 Engenharia de características

A modelagem de um evento de interesse é, em parte, uma tarefa subjetiva que requer compreensão do contexto, identificação e seleção de variáveis preditoras relevantes que sejam capazes de explicar este evento de maneira significativa. As primeiras referências de preditores que pudessem denotar comportamentos de risco nos contratos foram extraídas de trabalhos realizados por Sales *et al.* [47; 48]. Entretanto, são raros os trabalhos anteriores que descrevem detalhadamente as características utilizadas, seu cálculo ou sua importância na modelagem, possivelmente por seu caráter estratégico.

O trabalho desenvolvido por Sun *et al.* [52] estimou o risco de contratos na esfera federal (ainda que utilizando uma variável resposta relacionada a empresas). Com o objetivo de identificar características e fontes de dados de influência significativa na modelagem, este trabalho foi replicado em colaboração com seus autores. Por fim, com base na experiência de promotores e servidores do MPPB e outros órgãos da AP federal e estadual — por meio de debates, projetos conjuntos e da própria trajetória desta pesquisa — foram elaboradas características que pudessem, possivelmente, contribuir para o evento de rescisão contratual.

SAGRES

A partir do banco de dados do SAGRES, foram elaboradas 14 características sobre o histórico de participações, ganhos, abrangência e frequência de serviço, detalhes da licitação de origem e sobre o próprio contrato, e são descritas a seguir. Ainda, é relevante pontuar que todas as características foram calculadas até o momento de assinatura de cada contrato.

1. Modalidade da licitação que originou o contrato;
2. Valor total do contrato em reais (R\$);
3. Número de licitações nas quais a empresa participou;
4. Número de licitações nas quais a empresa foi vencedora;
5. Montante monetário acumulado pela empresa em licitações vencidas;

6. Percentual de vitórias da empresa em licitações;
7. Média de licitações nas quais a empresa foi participante por anos de atuação;
8. Média de licitações nas quais a empresa foi vencedora por anos de atuação;
9. Número de propostas enviadas pela empresa em licitações;
10. Média do número de propostas enviadas pela empresa em licitações por anos de atuação;
11. Total monetário em contratos firmados;
12. Razão entre o valor do contrato corrente e o valor acumulado em contratos pela empresa;
13. Número de contratos firmados pela empresa;
14. Média do número de contratos por ano.

Receita Federal

Com base na fonte de dados da RF, foram criadas quatro características adicionais, abrangendo aspectos cadastrais, composição societária, localização geográfica e o tempo de operação da empresa.

1. Número de sócios;
2. Porte econômico;
3. Categoria do logradouro onde está localizada (exemplo: avenida, praça, rodovia, rua, rural);
4. Tempo decorrido (em anos) entre a abertura da empresa e a data de início do contrato corrente.

As características acima foram descritas com o objetivo de viabilizar a replicação deste estudo ou servir de inspiração para futuros praticantes e não representam informações estratégicas de nenhum dos parceiros que contribuíram para a elaboração deste trabalho.

Por fim, um conjunto de 19 características foi gerado, abrangendo não apenas aspectos recentes dos contratos e empresas responsáveis, mas também seu histórico de participação, vitórias, proventos e assiduidade na prestação de serviços a AP.

6.1.2 Extração e tratamento

A inserção de dados em sistemas como o SAGRES é realizada manualmente pelos funcionários da AP em suas respectivas unidades gestoras, o que torna as informações suscetíveis a erros. Em vista disso, a remoção de registros possivelmente inconsistentes e contratos não relevantes ao escopo é detalhada a seguir. O arquivo de dados do banco foi fornecido originalmente em Oracle³⁹, o arquivo precisou ser convertido para um SGBD não proprietário (MariaDB 5.5.58⁴⁰), além da criação ou correção de índices de tabelas. Mais de 122 mil contratos foram assinados e incluídos no SAGRES entre 2014 e 2018 (com mais de 28 mil pessoas físicas ou jurídicas distintas contratadas)⁴¹.

Crítérios de filtragem

Com rescisões frequentes e alheias aos interesses deste estudo, contratos relacionados à contratação de profissionais de saúde (frequentemente médicos) foram removidos do conjunto de dados. Também foram desconsiderados contratos cuja pessoa jurídica responsável prestasse algum serviço essencial ao estado, tais como: companhia de fornecimento de energia (Energisa) ou água (CAGEPA), Caixa Econômica Federal, Banco do Brasil, Imprensa Nacional, Polícia Militar do Estado e outros 12 CNPJs⁴².

Na esfera federal, Sales *et al.* [48] consideram apenas contratos acima de um milhão de reais. Contudo, o calibre dos contratos das administrações municipais é, tipicamente, menor. Entre 2014 e 2018, por exemplo, 75% dos contratos têm valor inferior a R\$ 97.000. Contratos de valores exorbitantes ou extremamente baixos podem representar possível falha humana de preenchimento. No segundo caso, o custo de alocação de recursos para acompanhamento sequer compensaria seu próprio valor. Assim, foram removidos contratos de valor inferior a R\$ 800,00 e superior a R\$ 500.000.000,00.

³⁹ Oracle DB

⁴⁰ Maria DB

⁴¹ Dados parcialmente disponíveis no SAGRES Online e nos Dados abertos do TCE.

⁴² A lista completa de empresas pode ser vista no Apêndice C

Pela impossibilidade de cruzamento com outras etapas da contratação, contratos, licitações e empenhos com códigos de identificação (identificador próprio ou da empresa responsável) inválidos ou faltantes foram removidos. Contratos sem empenho associado, com código de licitação de origem ou data de homologação desconhecidos, e cujos códigos da unidade gestora ou descrições do objeto são desconhecidos, também foram removidos. Ainda, contratos com códigos identificadores faltantes ou inconsistentes também foram removidos. As características categóricas foram transformadas em variáveis binárias, como no caso do tipo de licitação de origem do contrato. Essa etapa de pré-processamento e filtragem resultou em um conjunto de contratos municipais com 17.138 registros.

6.1.3 Treino, teste e rótulos

A partir dos dados de mutações de interesse do Tramita, cada contrato foi rotulado como alto ou baixo risco: contratos previamente impedidos, paralisados, rescindidos, suspensos ou sustados foram considerados contratos de alto risco de rescisão. Caso contrário, foram rotulados como contratos de baixo risco de rescisão.

Com seleção aleatória, 80% do conjunto de dados foi utilizado para treinamento e validação cruzada do modelo, enquanto os 20% restantes foram destinados à avaliação da performance de generalização do modelo em uma amostra externa. Os arquivos de contratos de treino e teste possuem, respectivamente, divisões de 13.502/208 e 3.363/65 para as classes de baixo e alto risco.

6.1.4 Algoritmo e parametrização

Em determinados algoritmos de aprendizagem, nem todos os parâmetros para definição do modelo podem ser inferidos a partir do conjunto de dados durante o treinamento. Como visto na Seção 2.5, o XGBoost requer a definição de hiperparâmetros mínimos (ou uso de valores padrão das implementações propostas) e, uma vez que esses valores variam para cada conjunto de dados, são descritas abaixo as estratégias adotadas para sua definição.

É importante frisar que a busca pela combinação ideal de hiperparâmetros para otimizar o desempenho do modelo é, por si só, um campo de pesquisa. Ao considerar as diferenças significativas entre conjuntos de contratos em distintos contextos, este estudo não tem

por objetivo buscar exaustivamente uma configuração ótima de hiperparâmetros, mas sim explorar a modelagem do risco contratual dentro do escopo proposto.

A partir do método de amostragem hipercubo latino (HL), foram geradas faixas de valores semialeatórias para os hiperparâmetros. Os valores dos hiperparâmetros de modelos com performances promissoras foram manualmente selecionados e compõem a Tabela 6.1. Essa estratégia foi utilizada para fornecer uma direção inicial no conjunto de valores (com exceção do parâmetro de peso das classes de risco), evitando buscas exploratórias. Para evitar sobreajuste do modelo nos dados, foi utilizada a estratégia de CV com 3-folds. Demais hiperparâmetros foram mantidos em suas configurações de valores padrão⁴³.

Tabela 6.1: Hiperparâmetros - HL

Hiperparâmetro	Valores
n_estimators	[100; 200; 300; 400; 500; 600; 700; 800]
learning_rate	[0,012; 0,014; 0,016; 0,02; 0,025; 0,03; 0,035; 0,04; 0,045; 0,05; 0,055; 0,06; 0,07; 0,075; 0,08; 0,09; 0,1]
scale_pos_weight	[100; 120; 150; 200; 260; 300]
min_child_weight	[0,5; 0,8; 1; 2; 3; 4; 5; 6; 7; 10; 20]
reg_lambda	[0,1; 1; 4; 5; 8; 10; 13; 15; 20; 30; 50]

Tabela 6.2: Hiperparâmetros - *Grid Search*

Hiperparâmetro	Valores
n_estimators	[1600; 2000; 2100; 2200; 2300]
learning_rate	[0,005; 0,008; 0,01; 0,015; 0,020; 0,025; 0,035; 0,04]
scale_pos_weight	[100; 150; 200; 260]
reg_lambda	[1; 4; 8; 10; 13; 15]
min_child_weight	[0,7; 0,8; 1; 2; 4]

⁴³ Documentação XGBoost - Parametrização

A partir de rodadas de experimentação, de avaliação e de seleção manual, o conjunto de valores foi refinado, tendo como critério o aumento no valor da AUC-PR e a observação do erro na curva de aprendizado de cada modelo. O conjunto de hiperparâmetros utilizados para os treinamentos a seguir, por meio de *grid search*, pode ser verificado na Tabela [6.2](#).

6.2 Risco do contrato

Entre os 17.138 contratos municipais, 273 (distribuídos entre treino e teste) foram rotulados como de alto risco de rescisão, com base nos dados do Tramita. Essa proporção resulta em uma AUC-PR de 0,015 para um classificador aleatório hipotético treinado nesse conjunto de dados. Na avaliação da MR, a previsão de risco foi baseada no ranking do sumário das tipologias. Já nesta análise, no entanto, o risco de cada contrato é estimado por meio de um modelo XGBoost utilizando características do SAGRES e da RF, duas fontes de dados amplamente empregadas no acompanhamento e fiscalização de contratos, além dos hiperparâmetros apresentados na Tabela [6.2](#).

O modelo de melhor performance conta com os seguintes hiperparâmetros: *learning_rate* = 0.01; *min_child_weight* = 0,7; *n_estimators* = 1600; *reg_lambda* = 4 e *scale_pos_weight* = 260. A previsão do modelo para o conjunto de teste resultou em um intervalo de confiança para a AUC-PR entre 0,053 e 0,209; precisão = 0,09; revocação = 0,26; AUC-ROC = 0,60 e AUC-PR = 0,11.

As métricas de avaliação deste experimento não são diretamente comparáveis aos resultados da avaliação da MR, devido à significativa disparidade de volume e esparsidade entre os conjuntos de dados. No entanto, neste caso, os resultados apontam que, no pior caso – considerando o limite inferior do intervalo –, o modelo supera em sete vezes o classificador aleatório para este conjunto de dados. A Figura [6.1](#) ilustra o resultado da previsão de classificação do modelo para o conjunto de teste.

Para essa instância de previsão, 17 dos 65 contratos de alto risco foram corretamente identificados como verdadeiros positivos, enquanto 48 foram considerados falsamente como de baixo risco (falsos negativos). Na prática, pelo modelo, aproximadamente R\$ 12 milhões seriam devidamente acompanhados e fiscalizados, representando cerca de 7% do montante de alto risco de R\$ 196 milhões. Para os contratos de baixo risco, 3.192 foram corretamente

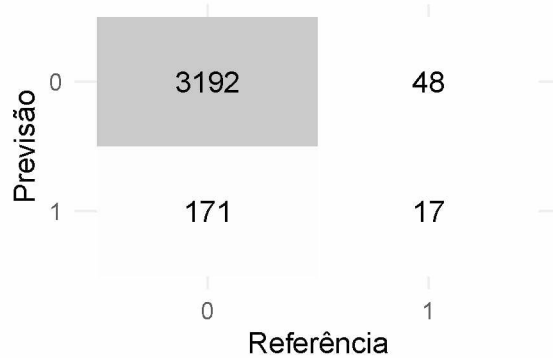


Figura 6.1: Classificação do conjunto de testes - Matriz de Confusão

classificados como verdadeiros negativos, totalizando cerca de R\$ 2,2 bilhões. Entretanto, 171 contratos, equivalentes a R\$ 112 milhões, foram erroneamente considerados como de alto risco pelo modelo (falsos positivos). Dos 188 contratos indicados pelo modelo como alto risco de rescisão, 17 resultariam em buscas frutíferas, enquanto os outros 171 poderiam ou não revelar possíveis irregularidades.

6.3 MR como filtro de contratos

Como visto na análise anterior, a modelagem do risco de contratos municipais é uma ferramenta promissora. A fim de verificar a capacidade de síntese de risco das tipologias e perspectiva de melhoria na performance do modelo – ao modelar o mesmo espaço de contratos passíveis de risco englobado pelas empresas da MR, o conjunto de contratos utilizado na seção anterior foi filtrado de forma a considerar apenas aqueles associados a empresas presentes na MR.

Dos 13.212 contratos de empresas da MR, 218 (1,6%) foram rotulados como alto risco de rescisão, o que define um valor base de 0,016 para um classificador aleatório considerando sua AUC-PR. Esse conjunto de dados foi separado para treinamento e teste com 80% e 20% das observações, com 10.433/168 e 2.561/50 observações nas classes de baixo e alto risco, respectivamente. Os mesmos conjuntos de hiperparâmetros e de preditores na aprendizagem (SAGRES e RF) do experimento anterior foram utilizados.

O modelo de melhor performance possui os seguintes hiperparâmetros: *learning_rate* = 0,008; *min_child_weight* = 4; *n_estimators* = 2000; *reg_lambda* = 15; *scale_pos_weight* =

100; e IC para a AUC-PR entre 0,070 e 0,277. Semelhantemente, no pior caso, o modelo supera o classificador aleatório hipotético desse conjunto de dados em quatro vezes. A matriz de confusão na Figura 6.2 informa o resultado para essa instância de previsão. A previsão do modelo resulta em uma precisão = 0,12, revocação = 0,22, AUC-PR = 0,15 e AUC-ROC = 0,59.

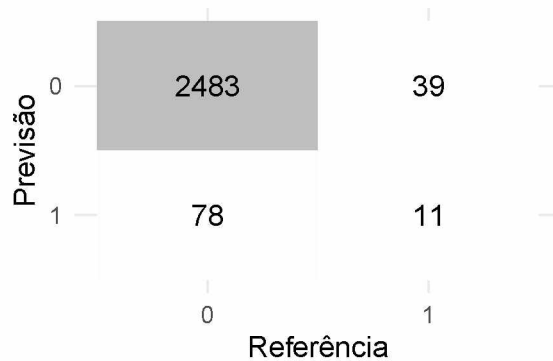


Figura 6.2: Classificação do conjunto de testes (contratos da MR) - Matriz de confusão

Como resultado, 11 dos 50 contratos de alto risco foram corretamente identificados pelo modelo, ao tempo que 39 foram falsamente classificados como baixo risco. Dos quase R\$ 181 milhões de reais referentes a contratos de alto risco de rescisão, apenas R\$ 6,7 milhões teriam o acompanhamento apropriado para melhor utilização de recursos. Os verdadeiros negativos representam R\$ 1,6 bilhão, enquanto falsos positivos pouco mais de R\$ 53 milhões.

6.4 Desbalanço entre classes de risco

Nas modelagens anteriores, a compensação do desbalanceamento entre as classes de risco no dado foi realizada através do aprendizado sensível ao custo, atribuindo pesos a cada classe. Com o objetivo de avaliar como estratégias de balanceamento teriam impacto na capacidade de aprendizagem do modelo, em especial no aprendizado da classe minoritária, dois modelos foram treinados utilizando a subamostragem (*undersampling*) e a sobreamostragem (*oversampling*) de dados. Tendo em vista a experimentação de outras estratégias de balanceamento, o hiperparâmetro *scale_pos_weight* foi suprimido durante o treinamento dos modelos desta seção.

O conjunto de contratos de treinamento é composto por 13.502 observações da classe

majoritária e 208 observações da classe minoritária, sendo que apenas o conjunto de treinamento foi artificialmente balanceado. As características do SAGRES e RF foram, novamente, utilizados como preditores de risco.

Para o *undersampling*, a classe minoritária foi mantida, e foram selecionados 208 registros aleatórios da classe majoritária para composição de um novo conjunto de treinamento com tamanho total de 416 contratos. Com 95% de confiança, o IC para o modelo *SAGRES RF Under* varia entre 0,026 e 0,090. Já para o *oversampling*, a classe minoritária foi replicada com reposição até atingir o número de registros da classe majoritária. Com 27.004 registros, a modelagem utilizando esse novo conjunto de treinamento resultou em um IC que varia entre 0,053 e 0,194.

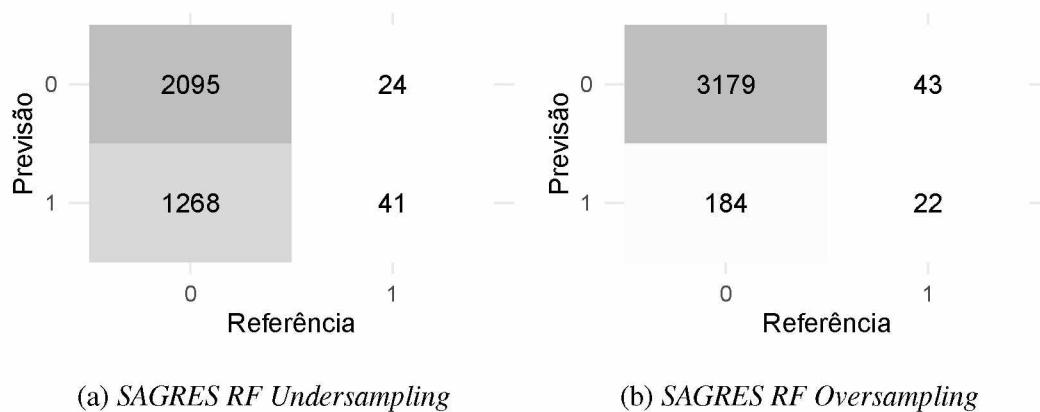


Figura 6.3: Comparação entre estratégias de mitigação do desbalanço - Matriz de confusão

Observa-se que, na estratégia de *undersampling* (Figura 6.3a), o modelo erra de forma mais uniforme na previsão de ambas as classes de risco. Apesar de um aumento nos verdadeiros positivos, essa abordagem resulta também em um aumento significativo no número de contratos classificados como falsos positivos – sete vezes maior ao comparar com o *SAGRES RF*. Como consequência de um modelo menos conservador, o número de falsos negativos é reduzido pela metade. Uma menor representatividade da classe negativa pode levar o modelo a identificar mais padrões na classe positiva, explicando o aumento dos falsos positivos.

Em contrapartida, ao replicar as observações da classe minoritária, o modelo treinado a partir da adoção do *oversampling* (Figura 6.3b) não apresenta mudanças significativas em seu comportamento. Ao passo que previu cinco contratos verdadeiros positivos a mais, 13 novas previsões falso-positivas também ocorreram. Embora a proporção entre as classes seja

equivalente, a mera replicação de contratos sabidamente problemáticos não demonstra, nesse contexto de modelagem, efetividade na melhoria de performance na classificação do risco.

A Figura 6.4 ilustra o comparativo de ICs das AUC-PR para os métodos de balanceamento artificial. Apesar da clara sobreposição entre os três intervalos, vê-se que o desempenho do modelo *SAGRES RF Under.* ainda, no melhor caso, é inferior aos demais. Por outro lado, o IC da diferença entre *SAGRES RF* e *SAGRES RF Over.* varia entre -0,092 e 0,113, o que demonstra ausência de diferença estatisticamente significativa entre ambas as estratégias de balanceamento entre classes – na prática, essas duas estratégias possuem efeitos equivalentes na previsão do modelo.

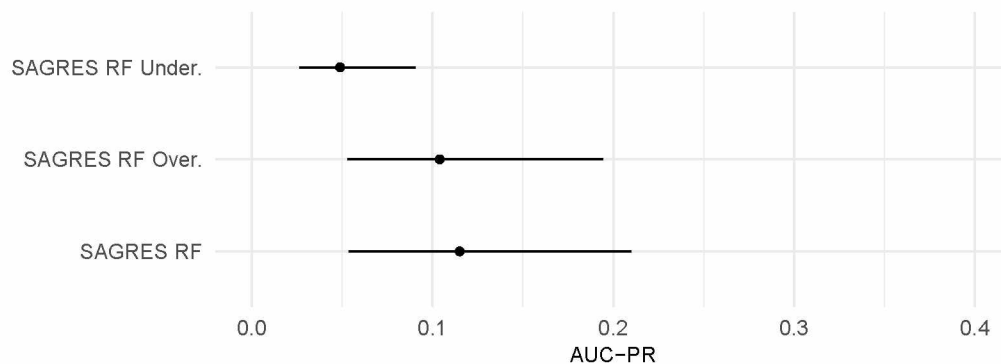


Figura 6.4: IC para 95% de confiança - Comparação entre desempenho de modelos

6.5 Discussão

Na Seção 4.3, abordou-se a necessidade do uso de métricas apropriadas para avaliação de modelos de risco em cenários de severo desbalanço entre classes. A AUC-PR foi utilizada para análise e comparação dos modelos propostos. Entretanto, estudos anteriores recorrem a outras métricas para avaliação de seus modelos. Assim, os comparativos abaixo foram realizados com base nas métricas disponíveis em cada estudo específico.

Neste capítulo, o esforço para estimação do risco de rescisão de contratos municipais resultou em dois modelos principais: um treinado com uma amostra ampla de contratos e outro considerando apenas contratos de empresas pertencentes a uma ferramenta de priorização de investigações em uso (MR). Com o primeiro modelo (*SAGRES RF*), o objetivo foi fornecer evidência para enquadrar o problema da previsão do risco de possíveis falhas no nível de

contrato por meio do dado de rescisão contratual. Os resultados apontam que, no cenário de pior caso, o preditor de risco tem desempenho sete vezes superior a um classificador aleatório considerando o valor de referência (AUC-PR) do próprio conjunto de dados.

O modelo *SAGRES RF* possui acurácia = 93%; precisão = 0,09; taxa de verdadeiros negativos (TNR) = 94%; AUC = 0,60 e $F_{0.5}$ = 0,1 e F_1 : 0,13. Em comparação, um modelo de redes neurais profundas treinado com um conjunto de contratos federais de grande porte [52] possui acurácia = 91%; precisão = 0,4; taxa de verdadeiros negativos (TNR) = 94%; AUC = 0,87; $F_{0.5}$ = 0,56 e F_1 = 0,48 (valor calculado a partir dos dados fornecidos). Com alta TNR, ambos os modelos erram menos na classificação de contratos de baixo risco (classe majoritária). Entretanto, o *DNN* possui maior capacidade de identificação de contratos de alto risco de fato (verdadeiros positivos). É provável que a diferença significativa na capacidade de identificação de contratos de alto risco entre os modelos se dê pela disparidade na proporção de contratos de alto risco no próprio conjunto de dados – o conjunto utilizado por Sun *et al.* possui sete vezes mais observações da classe minoritária do que o utilizado no treinamento do modelo *SAGRES RF*.

Gallego *et al.* [27] propõem modelos similares aos apresentados nesta análise e também utilizam um algoritmo baseado em árvore (Gradieng Boosting Machine – GBM). Contudo, estimam o risco de contratos colombianos em relação a três indicadores de irregularidades distintos. O modelo que utiliza a indicação de risco de ineficiência contratual como rótulo possui a melhor performance para a AUC-PR. Apesar das métricas de avaliação para todos os modelos gerados no trabalho não estarem explicitamente disponíveis, é possível realizar uma comparação visual de suas curvas de precisão-revocação. Na Figura 6.5, vê-se que a AUC-PR do modelo para os contratos colombianos é superior à do *SAGRES RF*. Ainda, nesse estudo, para qualquer uma das três variáveis resposta experimentadas, o cenário de desbalanceamento é menos severo do que o enfrentado no conjunto de contratos municipais da Paraíba. A classe minoritária para o rótulo do melhor modelo representava 11% dos dados totais.

Nos dois estudos citados acima, a classe de alto risco na variável resposta (classe de interesse na modelagem) tem proporções consideravelmente maiores do que a utilizada nessa análise. No entanto, enquanto, no primeiro caso, essa variável resposta faz referência ao histórico de sanções das empresas responsáveis pelos contratos, no segundo caso ela denota

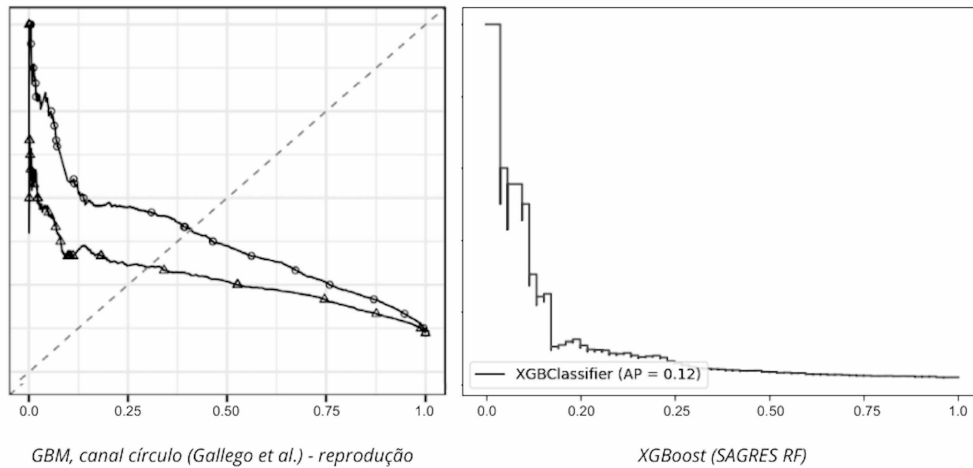


Figura 6.5: Comparativo de curva de precisão-revocação - GBM vs. XGBoost

a possibilidade de irregularidades. À vista disso, se o aprendizado é essencialmente a identificação de padrões que relacionem descritores à variável resposta, isso significa dizer que, nos casos mencionados, há mais exemplos nos quais o modelo pôde aprender os padrões do comportamento de risco para cada cenário.

Apesar do paralelo entre as métricas dos modelos, é imprescindível ressaltar que os conjuntos de dados utilizados para rotulagem dos diversos grupos de contratos podem possuir entidade-referência, contexto e semântica distintas. Todavia, embora denotem um tipo de risco em alguma medida, não quantificam exatamente os mesmos comportamentos. No segundo estudo comparado [27], em especial, as variáveis respostas no grão de contrato denotam apenas mera indicação de risco (uma tipologia de risco, em outras palavras), e não contratos previamente impedidos ou interrompidos, de fato – o que contribui diretamente para um conjunto de rótulos mais numeroso.

Já o segundo modelo proposto, *SAGRES RF (contratos MR)*, propôs simular o universo de contratos passíveis de fiscalização das empresas presentes na MR para verificar a capacidade de identificação/filtragem do risco nas tipologias propostas pela ferramenta. Comparativamente, esse modelo possui maior precisão, menor revocação e maior AUC-PR do que o *SAGRES RF*. Na prática, isso significa que o *SAGRES RF (contratos MR)* prevê menos contratos falsos positivos ao custo de mais contratos falsos negativos – o modelo é mais conservador e tem a tendência a classificar um número maior de contratos como baixo risco. No contexto de contratação, consequentemente, o montante financeiro passível de fiscalização

é menor, assim como as chances de investigar contratos possivelmente não irregulares ou falhos também.

Entre *SAGRES RF* e *SAGRES RF (contratos MR)*, considerando os respectivos conjuntos de teste, há: uma redução no percentual de contratos classificados como falsos positivos (de 4,98% para 2,98%), um aumento na proporção de falsos negativos (de 1,4% para 1,9%), enquanto o percentual de verdadeiros positivos decai (de 0,49% para 0,42%). A fim de selecionar um modelo para continuação das demais análises, a diferença entre ambos foi calculada por meio de um IC para sua AUC-PR.

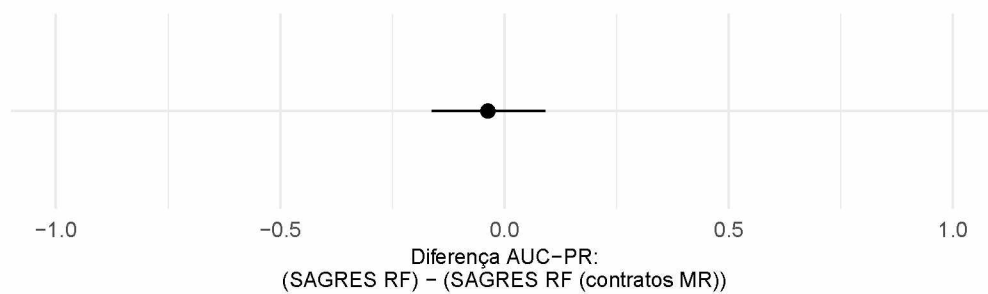


Figura 6.6: IC da diferença entre os modelos *SAGRES RF* e *SAGRES RF (contratos MR)*

Considerando as previsões, duas amostras independentes, foi calculado o IC da diferença entre suas AUC-PR, ilustrado na Figura 6.6, que varia entre -0,162 e 0,091. Com uma confiança de 95%, o IC mostra a faixa de valores plausíveis e, nesse caso, indica que é provável que não haja diferença estatisticamente significativa entre o desempenho de ambos os classificadores. Mesmo que, em alguma medida prática, o filtro proporcionado pela MR possa selecionar contratos de empresas mais propensas ao risco para modelagem, não é possível afirmar que o modelo treinado com este conjunto de dados possui desempenho superior. Assim, o conjunto de contratos sem filtro será utilizado para subseqüentes modelagens e análises.

Na modelagem de risco de rescisão de contratos municipais, não há diferença significativa entre as estratégias de balanceamento – assim como apontado por Gallego *et al.* [27]. Se, por um lado, a sub-representação da classe negativa pode enviesar o aprendizado em relação à classe positiva, a mera replicação das observações não produz padrões relevantes que auxiliem na identificação mais eficiente da classe minoritária. Esse comportamento corrobora resultados anteriores [41] de que estratégias de balanceamento podem não ter efeito prático

relevante para aplicação em conjuntos de dados com eventos reais.

De modo geral, até aqui, tomando em consideração a ineficiência da estimação em nível de empresa e os resultados promissores, ao comparar o AUC-PR de cada modelo proposto com um classificador aleatório para cada conjunto de dados, pode-se dizer que há evidência para abordar o problema de estimação de risco contratual no nível da entidade de contrato. Contudo, não há evidência de que utilizar estratégias para balanceamento artificial das classes de risco tenha impacto positivo na eficácia da previsão.

Limitações

Uma primeira limitação dos experimentos e análises realizadas neste capítulo relaciona-se ao tamanho da amostra de contratos de alto risco utilizada para rotulagem do conjunto de contratos empregado para treino e teste. Essa amostra é, por si, subnotificada e sub-representada.

Ainda, como consequência de mecanismos de controle e fiscalização tipicamente reativos, um viés de seleção inerente à natureza do exercício do controle pode afetar a composição desse conjunto de rótulos. Uma vez que os contratos utilizados como rótulos foram selecionados por agentes humanos, os padrões de risco passíveis de aprendizado pelos modelos podem estar enviesados por características comuns na tomada de decisão humana.

Por fim, a replicação desse estudo, para comparação semanticamente consistente de resultados, está condicionada à existência de um conjunto de rótulos semanticamente similar – o que pode não estar disponível, ao menos publicamente, em todos os estados brasileiros.

Capítulo 7

Experimentação de outros preditores de risco

No capítulo anterior, foram propostos e avaliados modelos de previsão de risco de rescisão para contratos municipais utilizando dois conjuntos de características frequentemente empregados no acompanhamento de contratações. Esses conjuntos condensam informações sobre as duas principais entidades envolvidas nesse tipo de modelagem: os próprios contratos e as empresas responsáveis por eles. Neste capítulo, analisa-se o impacto da inclusão de características relacionadas ao tema do contrato e a um fator econômico municipal. Além disso, avalia-se o impacto na performance do modelo no cenário da ausência de ao menos uma fonte de dado, e da influência do valor do contrato no aprendizado – tendo em vista o custo do erro de classificação de contratos financeiramente mais relevantes.

7.1 Objeto do contrato como característica

Na literatura sobre previsão de risco em contratos públicos, não é comum o detalhamento das características preditoras utilizadas para treinamento dos modelos de AM, com exceção de poucos trabalhos menos recentes. Em geral, os praticantes limitam-se a descrever brevemente seus contextos de abrangência ou fonte de dado de origem. Em específico na estimação do risco ⁴⁴ de contratos, as características utilizadas foram, a partir das informações disponíveis em cada estudo prévio disponível e considerado neste trabalho (detalhadas

⁴⁴A tipificação do risco pode variar a cada estudo.

na Seção 3.1.2), dados numéricos ou categóricos relativos ao contrato em si ou à empresa responsável.

Com base na experiência prática de especialistas – dois promotores e três servidores públicos –, contratações em determinadas áreas consistentemente apresentam falhas, irregularidades ou grandes desafios de gestão. Obras, fornecimento de combustível, coleta de lixo, determinados tipos de aquisições de alimentos e prestações de serviços são exemplos de temas que tipicamente saltam aos olhos no quesito suspeição de risco em relação ao seu planejamento ou aos contratos, não apenas no contexto dos municípios paraibanos.

A partir desse pressuposto, utilizar o tema/área do contrato para modelagem de risco de rescisão poderia resultar em melhora na performance do modelo, caso o conjunto de contratos sabidamente problemáticos possua associação frequente com determinada área de contratação. É possível, ainda, pensar em uma modelagem individual para cada área de contratação. Contudo, o cenário de desbalanceamento entre classes seria ainda mais agravado, tendo em vista a quantidade de registros disponível para a variável resposta. Ainda, a depender do conjunto de dados utilizado, é possível que ao menos uma área não tivesse nenhum contrato rotulado como alto risco de rescisão para modelagem.

Para contratos da esfera federal, é possível identificar minimamente a área à qual um contrato está associado por meio da natureza de despesa (função e subfunção orçamentárias) do gasto em sua execução orçamentária. Para os municípios paraibanos esse campo de livre preenchimento está disponível; entretanto, é frequentemente incompleto, inconsistente ou não preenchido. Por esse motivo, o campo da descrição do objeto contratual será utilizado para inferir seu tema de abrangência.

A Tabela 7.1 exibe exemplos de descrição de objeto extraídas do conjunto de dados. Pelo livre preenchimento do campo, vê-se que o detalhamento e a consistência do preenchimento da descrição do objetivo ficam a cargo do agente humano responsável. Além disso, algumas descrições podem apresentar ambiguidade nos termos temáticos, como, por exemplo, o primeiro e o sétimo objetos descritos na tabela, em que são adquiridos combustíveis e peças automotivas, respectivamente, porém destinadas aos veículos operantes na área da saúde.

A descrição do objeto do contrato foi utilizada com sucesso para categorização de gastos [23] e comparação de preços de itens de compra com descrições semelhantes [13]. Entretanto, na literatura atual, não há referências sobre o uso ou avaliação de impacto de atributos

Nº	Descrição do objeto
1	"Aquisição parcelada de Combustíveis, derivados de petróleo, lubrificantes e filtros, destinados a frota de veículos de propriedade da Prefeitura Municipal e Fundo Municipal de Saúde, sejam contratados, locados à disposição ou vinculados a atividade pública do município (...)"
2	"Fornecimento parcelado de bolsa de colostomia e urostomia, para atender as necessidades da Secretaria de Saúde do Município (...)"
3	"Contratação de empresa especializada, cujo critério de seleção da proposta mais vantajosa será a de menor preço global, para a execução de serviço de pavimentação em paralelepípedo junto ao município (...)"
4	"Aquisição de Gêneros Alimentícios para Merenda dos Alunos da Rede Municipal de Ensino e para o desenvolvimento das Ações, Atividades e Programas de todas as Secretarias da Prefeitura do Município (...)"
5	"AQUISIÇÃO DE PEIXE"
6	"Prestar o fornecimento parcelado de gases medicinais para atender as necessidades do Hospital Materno Infantil (...)"
7	"Fornecimento de Peças Automotivas destinadas a manutenções corretivas e preventivas de veículos pesados, pertencentes e/ou locados a esta Edilidade e ao Fundo Municipal de Saúde."

Tabela 7.1: Exemplos de descrições de objetos contratuais

temáticos relacionados ao objeto adquirido como preditores na estimação do risco de rescisão de contratos. Esta análise visa preencher essa lacuna e servir como ponto de partida para trabalhos futuros.

7.1.1 Extração de tópicos

A técnica do LDA (detalhada na Seção [2.4.1](#)) considera que cada documento é composto por proporções de tópicos (temas) e que cada tópico é composto por proporções de palavras relevantes. As estruturas semânticas latentes presentes na descrição de cada produto ou serviço (temas) são capturadas através do uso dessa técnica. Uma vez definida a quantidade de

tópicos que se deseja identificar em uma coleção de documentos, o algoritmo é responsável por rearranjar a proporção da ocorrência de tópicos para cada documento e de palavras para cada tópico.

Tal qual outros algoritmos de processamento de linguagem natural, o LDA também se baseia na frequência de ocorrência das palavras para construção de seu vocabulário. Assim, as descrições dos contratos foram fragmentadas em palavras, sendo removidos cerca de 900 termos irrelevantes à caracterização do objeto (*stopwords*, disponíveis no Apêndice [D](#)), e termos com um ou dois caracteres alfanuméricos. Também foram ignorados termos que ocorrem em número inferior a dez vezes, bem como os que ocorrem em mais de 35% dos documentos.

Para geração do corpus, unigramas e bigramas (sequências de uma ou duas palavras, respectivamente) foram considerados. O número de tópicos foi definido a partir de uma abordagem incremental e iterativa com avaliações manuais, levando em consideração a coerência. Foram consideradas tanto a coerência semântica entre os conjuntos de palavras associadas quanto a distribuição do número de contratos mais fortemente vinculados a cada tópico. Quanto maior o número de tópicos, maiores as chances de que contratos de um mesmo tema sejam distribuídos entre diferentes tópicos. Por outro lado, com um número reduzido de tópicos, contratos de temas distintos tendem a ser agrupados em um mesmo tópico. O objetivo, então, é alcançar um equilíbrio que permita a geração de tópicos semanticamente coesos e com significados idealmente distintos entre si.

Nessa perspectiva, conjuntos de 10, 15, 20, 25, 30 e 35 tópicos foram experimentados. Por fim, foram considerados 20 tópicos para o treinamento do modelo. O algoritmo iterou sob o *corpus* por 50 vezes, e cada palavra foi associada a pelo menos um tópico. A partir dos conjuntos de dados utilizados para treino e teste no Capítulo [6](#) - Seção [6.2](#), o modelo LDA foi treinado com as descrições do conjunto de treinamento e aplicado para definição das probabilidades contrato-tópicos no conjunto de testes.

Para cada contrato, foi gerada uma distribuição de probabilidade que indica sua associação com cada um dos 20 tópicos. Essa matriz reflete a probabilidade de cada contrato estar mais ou menos relacionado a determinado tópico, abrangendo aspectos intangíveis do contrato. As probabilidades contrato-tópico (a partir de agora referidos apenas como Tópicos) foram utilizadas como características preditivas para treinamento de modelos de risco

de rescisão contratual. Assim, a Figura 7.1 ilustra o processo.

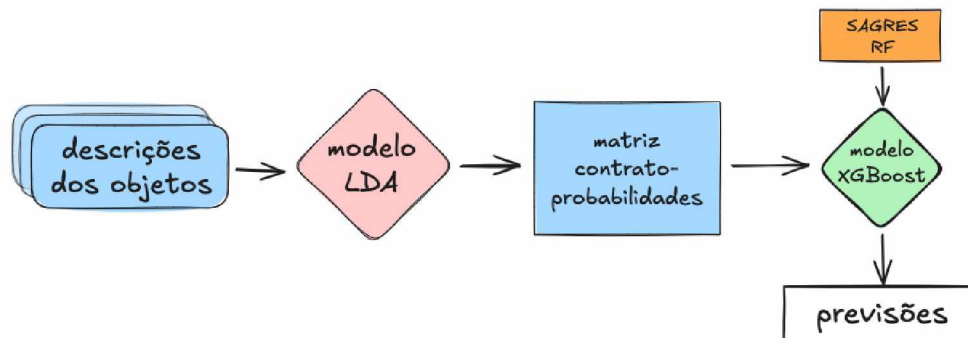


Figura 7.1: Ilustração da geração de tópicos por contrato e treinamento do modelo XGBoost

7.1.2 Resultados

Dando continuidade à modelagem do risco de rescisão de contratos, expandiu-se o conjunto de características preditoras com a inclusão dos Tópicos aos dois conjuntos já utilizados (SAGRES e RF, 6.2). Dessa forma, o modelo *SAGRES RF Tópicos* resultou em um intervalo de confiança de 95% para a AUC-PR na previsão do conjunto de dados de teste variando entre 0,057 e 0,213.

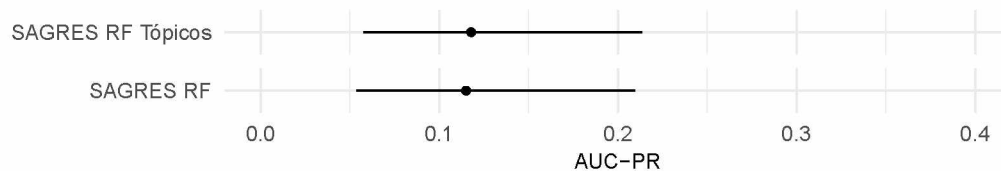


Figura 7.2: IC para 95% de confiança - Comparação entre desempenho de modelos

Embora haja uma melhoria marginal no desempenho do modelo comparado ao *SAGRES RF* (IC [0,0538; 0,209]), a sobreposição dos intervalos sugere que o conjunto de características que denota a área de atuação do contrato não imprime diferença estatisticamente significativa na previsão de risco do modelo. O IC da diferença entre *SAGRES RF Tópicos* e *SAGRES RF* varia entre [-0,106; 0,107] e corrobora essa interpretação. A Figura 7.2 ilustra a comparação dos ICs.

A Figura 7.3 exibe a matriz de confusão com erros e acertos para essa instância de previsão da classificação. Embora o desempenho geral dos modelos seja semelhante – ou seja,

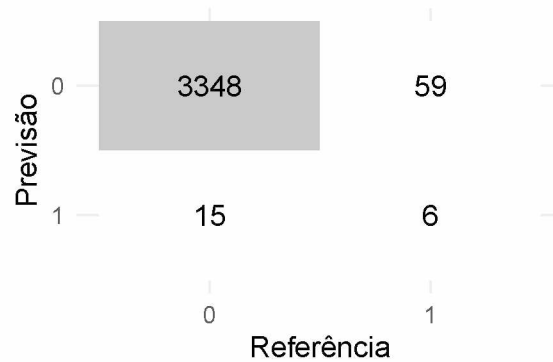


Figura 7.3: Classificação do conjunto de testes - Matriz de confusão

a capacidade de distinguir entre as classes de risco não tenha mudado significativamente, precisão e revocação seguem comparáveis –, a matriz de confusão revela que a inclusão do conjunto de Tópicos torna o modelo mais conservador em suas classificações. Entre *SAGRES RF* e *SAGRES RF Tópicos*, há diminuição significativa do número de falsos positivos (de 171 para 15), diminuição dos verdadeiros positivos (de 17 para 6) e aumento do número de falsos negativos (de 48 para 59).

Na prática, o modelo *SAGRES RF Tópicos* produz menos "alarmes falsos" em relação aos contratos possivelmente problemáticos, ao custo de identificar um número menor de contratos sabidamente problemáticos. A previsão do modelo resulta em uma precisão = 0,28; revocação = 0,09; AUC-PR = 0,11 e AUC-ROC = 0,54. Os contratos verdadeiros negativos totalizam R\$ 2,3 bilhões, falsos positivos R\$ 16 milhões, falsos negativos R\$ 190 milhões e verdadeiros positivos R\$ 6,2 milhões (valores aproximados).

7.2 Característica econômica dos municípios

Os conjuntos de características utilizados nas modelagens de risco anteriores forneceram informações históricas sobre os próprios contratos, as empresas responsáveis e o seu histórico de participação em licitações. Na etapa de qualificação deste trabalho de tese, uma auditora de contas públicas da Controladoria Geral do Estado da Paraíba (CGE-PB) levantou a hipótese de que municípios com economias mais produtivas tendem a firmar um maior número de contratos públicos, o que poderia levar, conseqüentemente, a um maior número de contratos com possíveis falhas ou irregularidades.

Um estudo em municípios húngaros avaliou a relação entre risco de corrupção e qualidade da competitividade entre licitantes, utilizando a renda média tributável *per capita* como um fator relevante [53]. Os resultados indicaram que cidades mais ricas apresentam menor risco de corrupção em compras públicas e maior controle sobre irregularidades. Dessa forma, com base na sugestão da especialista e nesse resultado da literatura, considerou-se que diferenças econômicas entre os municípios, refletidas no porte de suas economias, poderiam ser características preditoras relevantes para o risco de rescisão contratual.

O Produto Interno Bruto (PIB), que representa o valor monetário total de todos os bens e serviços produzidos em um determinado período, é uma métrica amplamente utilizada para descrever o tamanho de uma economia. O PIB *per capita*, por sua vez, reflete o valor produzido pelo município em relação à sua população⁴⁵. Os dados públicos e abertos do Instituto Brasileiro de Geografia e Estatística (IBGE) disponibilizam o PIB *per capita* dos municípios paraibanos para o ano de 2018. Assim, juntamente ao SAGRES e RF, o PIB *per capita* de cada município foi utilizado como preditor de risco na modelagem.

7.2.1 Resultados

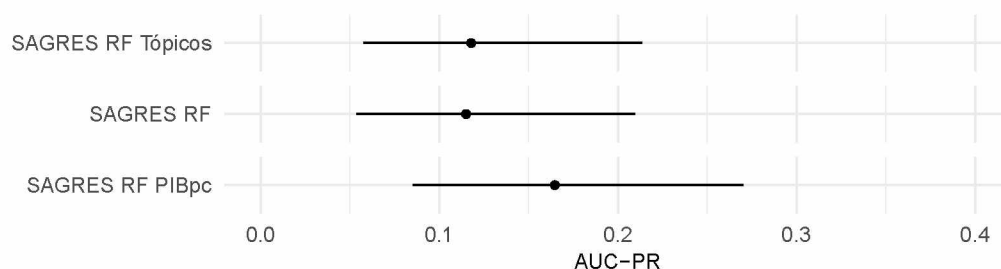


Figura 7.4: IC para 95% de confiança – Comparação entre desempenho de modelos

Com 95% de confiança, o IC para AUC-PR do modelo *SAGRES RF PIBpc* varia entre 0,085 e 0,270. A Figura 7.4 compara este IC com os de modelos anteriores. Ainda que sobrepostos os intervalos, é perceptível uma possível melhoria de performance do classificador ao adicionar o PIB *per capita* dos municípios como preditor. Ao calcular o IC da diferença ([−0,071; 0,171]) entre *SAGRES RF PIBpc* e *SAGRES RF* (exibido na Figura 7.5), vê-se que não há diferença estatisticamente significativa entre ambos os modelos – ainda que seja plau-

⁴⁵ Produto Interno Bruto (PIB) - IBGE

sível que o modelo *SAGRES RF PIBpc* tenha uma performance superior, quando comparada ao seu concorrente.

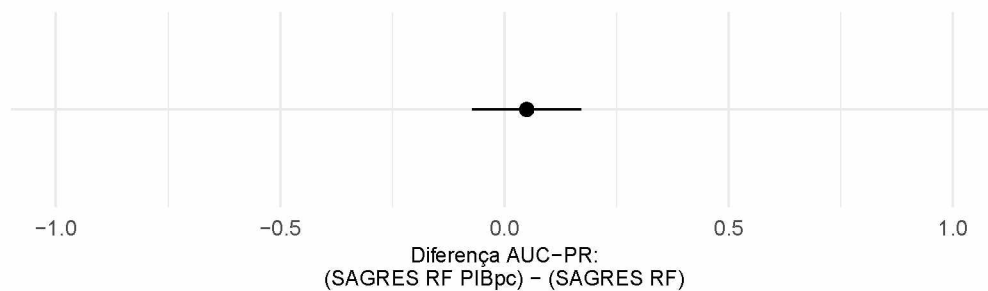


Figura 7.5: IC da diferença entre os modelos *SAGRES RF PIBpc* e *SAGRES RF*

A Figura 7.6 exibe a previsão do modelo para o conjunto de testes. Essa instância de previsão demonstra que o uso do *PIBper capita* torna o modelo mais conservador, principalmente em relação à classificação de falsos positivos, ao comparar com *SAGRES RF*. Há diminuição significativa de falsos positivos (de 171 para 102), aumento marginal dos falsos negativos (de 48 para 49) e diminuição também marginal dos verdadeiros positivos (de 17 para 16). Assim como no experimento anterior, o modelo classificaria menos contratos falsamente arriscados; porém, neste caso, sem que sua capacidade de identificar contratos verdadeiramente problemáticos seja comprometida na mesma proporção.

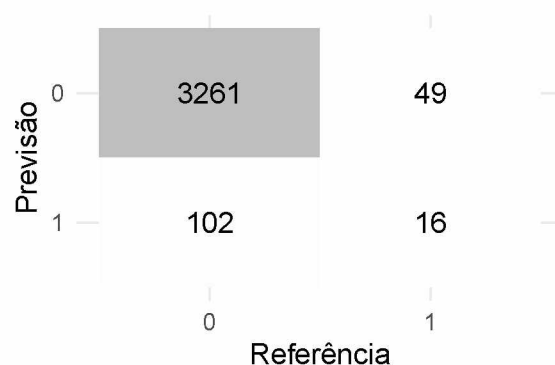


Figura 7.6: Classificação do conjunto de testes – Matriz de confusão

A previsão do modelo resulta em uma precisão = 0,13; revocação = 0,24; AUC-PR = 0,16 e AUC-ROC = 0,60. Os contratos verdadeiros negativos totalizam R\$ 2,2 bilhões, falsos positivos R\$ 56 milhões, falsos negativos R\$ 122 milhões e verdadeiros positivos R\$ 74 milhões (valores aproximados).

7.3 Pesos a partir do valor do contrato

As tarefas de acompanhamento e fiscalização dos contratos públicos envolvem a alocação de recursos humanos e financeiros para sua execução. O próprio processo de controle consome recursos públicos [21]. Diante disso, a AP deve analisar a relação custo-benefício entre o esforço empregado no acompanhamento da execução do contrato e o valor do contrato em si ou o benefício que pode agregar à sociedade. Contratos em áreas críticas ou de grande valor, por exemplo, poderiam receber maior atenção.

Na estimação do risco de rescisão contratual, errar a classificação de risco de um contrato de valor alto tem um maior custo à AP do que errar a classificação de um contrato de preço irrisório. Nesse contexto, durante o processo de aprendizagem do modelo, é possível penalizar casos em que a classificação de contratos mais “caros” seja feita de maneira incorreta. Para isso, uma lista de pesos foi calculada a partir da função de normalização min-max (detalhada na Seção 4.4) e aplicada no treinamento de modelo, de modo que ele minimize o erro geral considerando o peso de cada contrato.

7.3.1 Resultados

O modelo treinado utilizando a especificação de pesos para cada contrato a partir do seu valor (*SAGRES RF PIBcp (pesos obs.)*) resultou em um IC com 95% de confiança para a AUC-PR entre 0,072 e 0,261, com precisão = 0,119; revocação = 0,307; AUC-PR = 0,153 e AUC-ROC = 0,631. De modo geral, vê-se que não há melhora no desempenho do classificador ao adotar a estratégia de pesos relacionados aos valores dos contratos em relação ao modelo *SAGRES RF PIBpc*. O IC da diferença entre eles corrobora essa percepção – variando entre -0,149 e 0,120. O IC sugere que não há diferença estatisticamente significativa entre os modelos.

Entretanto, ao analisar a previsão individual e os valores dos contratos classificados como alto e baixo risco de rescisão, tem-se: R\$ 2,2 bilhões em contratos verdadeiros negativos (3.215), R\$ 85 milhões em falsos positivos (148), R\$ 111 milhões em falsos negativos (45) e R\$ 84 milhões em verdadeiros positivos (20). A Figura 7.7 ilustra a matriz de confusão para essa instância da previsão no conjunto de dados de testes.

Comparativamente, através das previsões dos modelos, com o *SAGRES RF PIBcp (pesos*

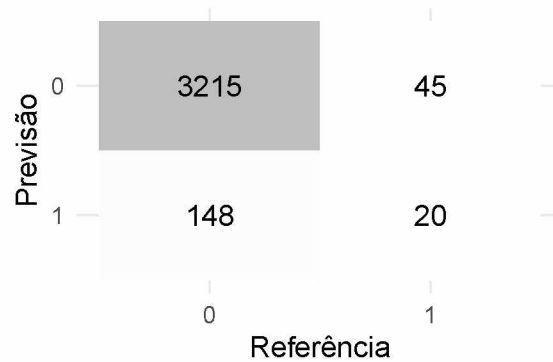


Figura 7.7: Classificação do conjunto de testes – Matriz de confusão

obs.) seria possível acompanhar apenas quatro contratos, de fato, de alto risco, a mais, em relação ao *SAGRES RF PIBpc*, porém o montante em recursos acompanhado seria superior em mais de R\$ 10 milhões. Por outro lado, o custo disso seria dispendioso para acompanhamento de 46 contratos falsos positivos a mais (uma diferença de mais de R\$ 29 milhões).

7.4 Impacto de diferentes conjuntos de características

O acesso, a disponibilidade, a consistência e a integridade de fontes de dados que possam aperfeiçoar o trabalho diário de agências governamentais e iniciativas cívicas de controle são desafios comuns e frequentes. No que tange aos processos de contratação e execução orçamentária, especificamente, alguns desses bancos de dados podem estar publicamente acessíveis – em formato aberto ou não ⁴⁶, enquanto outros são de difícil obtenção e/ou uso.

Os dados da RF, por exemplo, apenas recentemente tornaram-se massivamente disponíveis ao público para descarga e consulta (sendo ainda necessário cruzar diversos arquivos para obter todas as informações acerca de um CNPJ), através do portal de dados abertos do governo federal. Antes disso, havia a necessidade de consulta individualizada por CNPJ, com a obrigatoriedade do preenchimento de um CAPTCHA para acesso às informações. À época, casos de cobranças financeiras para usuários com interesse no acesso à base completa do dado público foram noticiados ⁴⁷.

⁴⁶O que são dados abertos? - Open Knowledge Foundation

⁴⁷"Serpro continua cobrando por acesso aos dados públicos sem amparo legal – Capital Digital."

Os dados abertos do TCE-PB, responsável pelo SAGRES-PB, referentes ao estado da Paraíba e aos seus municípios, não contemplam todas as etapas do processo de contratação e execução orçamentária⁴⁸. Além disso, dados equivalentes aos contidos no SAGRES podem não estar publicamente disponíveis em outros estados brasileiros. Também é importante ressaltar que, em muitos países, o acesso a dados públicos estruturados e confiáveis ainda é precário.

Com isso, é relevante quantificar o impacto de cada conjunto de características provenientes das fontes de dados utilizadas no desempenho preditivo dos modelos de AM. Dessa maneira, agências governamentais, praticantes acadêmicos futuros ou organizações da sociedade civil podem basear-se nessa análise para replicação do método e estimação do risco de rescisão em seus próprios conjuntos de contratos.

Considerando os conjuntos de características SAGRES, RF e Tópicos, são comparados os desempenhos de modelos em cenários nos quais ao menos um desses conjuntos não está disponível. A Figura 7.8 exibe os ICs com 95% de confiança para a AUC-PR de cada modelagem da combinação dos conjuntos de características. O IC do modelo *SAGRES RF PIBpc* foi incluído na imagem para fins de comparação.

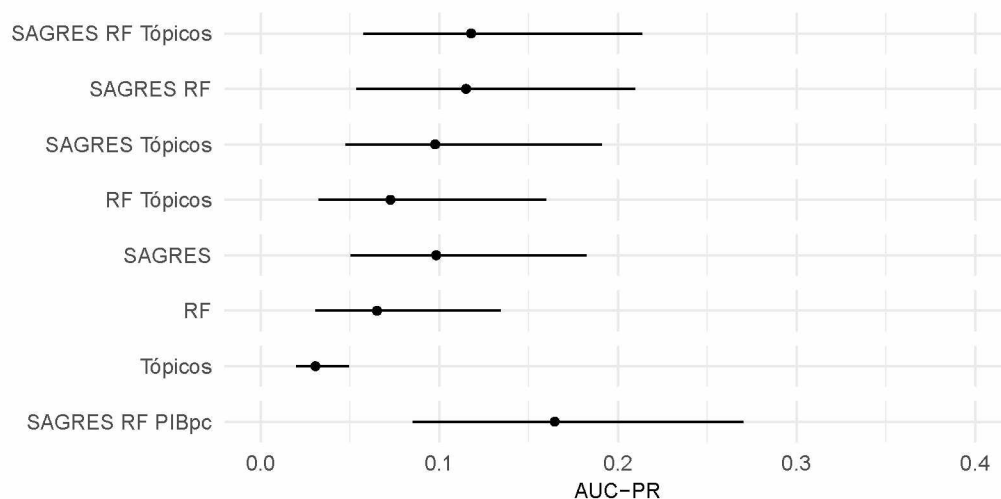


Figura 7.8: Impacto da indisponibilidade de conjuntos de características – IC 95% de confiança para AUC-PR

O limiar inferior dos ICs de todos os modelos apresentados está acima do desempenho de um classificador aleatório para este conjunto de dados ($AUC-PR = 0,015$), indicando que

⁴⁸Dados Abertos – TCE-PB.

os modelos possuem capacidade preditiva mínima para distinção entre ambas as classes de risco. Ao considerar cenários com apenas um conjunto de características para treinamento dos modelos, o conjunto Tópicos produz a separação menos efetiva entre as classes de baixo e alto risco de rescisão, gerando o modelo mais próximo de um palpite aleatório (IC = [0,020; 0,049]).

A ausência das características do SAGRES pode afetar negativamente a capacidade preditiva de todos os modelos desse cenário (IC *RF Tópicos* = [0,032; 0,159]; IC *RF* = [0,030; 0,134], IC *Tópicos* = [0,020; 0,049]). Em contraste, a ausência das características da RF e Tópicos tem efeito marginal nas performances (IC *SAGRES Tópicos* = [0,047; 0,191], *SAGRES RF* = [0,053; 0,209], *SAGRES* = [0,050; 0,182]). Essa observação ressalta a importância de dados históricos do fornecedor ou do próprio contrato para a efetividade da modelagem do risco de rescisão contratual por meio de AM. Por fim, o classificador *SAGRES RF Tópicos* alcança os valores mais altos em relação aos limites inferiores e superiores dos ICs ([0,057; 0,213]).

O IC da diferença entre *SAGRES RF Tópicos* e *Tópicos* varia entre -0,0009 e 0,154, sugerindo que o desempenho do primeiro modelo é praticamente equivalente ou significativamente superior à do segundo.

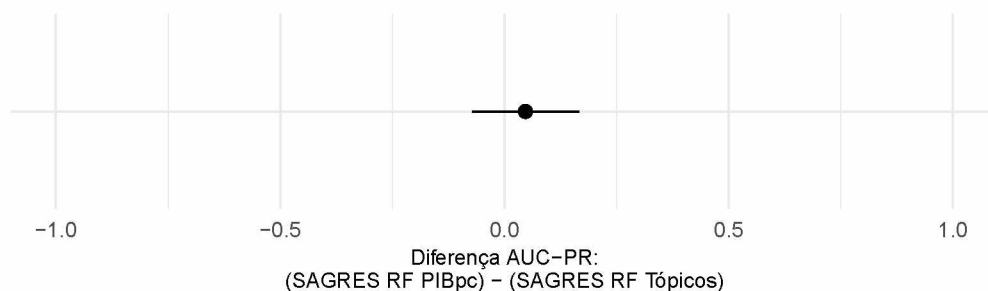


Figura 7.9: IC da diferença entre os modelos *SAGRES RF PIBpc* e *SAGRES RF Tópicos*

A Figura 7.9 exibe o intervalo de confiança da diferença entre o *SAGRES RF PIBpc* e o modelo mais promissor nesse cenário de comparação (*SAGRES RF Tópicos*), considerando o limite superior dos intervalos de confiança. O IC para AUC-PR varia entre -0,071 e 0,166, indicando que é plausível que o primeiro modelo tenha um desempenho superior ao segundo, embora não haja diferença estatisticamente significativa entre ambos no nível de 95% de confiança.

7.5 Discussão

No Capítulo 6, a eficácia dos modelos de AM apresentados foi diretamente comparada à de trabalhos anteriores (ainda que desenvolvidos para governos, esferas de governo e com variáveis resposta distintas) para validar o enquadramento da estimação do risco de rescisão contratual no nível de contrato, utilizando rótulos específicos sobre os próprios contratos. Neste capítulo, foram avaliados experimentalmente os impactos do uso de novas características e da estratégia de atribuição de pesos de acordo com o valor do contrato para a previsão do modelo. A partir disso, vê-se que a modelagem do risco de rescisão contratual não é uma tarefa trivial, cuja análise de resultados e direcionamento necessita ponderações financeiras, administrativas e de governança. A seguir, o modelo *SAGRES RF*, apresentado na Seção 6.2, é utilizado como ponto de partida para comparações.

O uso de características semanticamente latentes, extraídas da descrição do objeto do contrato como preditoras do modelo, não resultou em uma melhora significativa na performance preditiva geral do modelo (AUC-PR). Considerando o limite superior dos ICs (melhor caso), o modelo *SAGRES RF Tópicos* é apenas 5% superior ao *SAGRES RF* ao comparar os valores para a AUC-PR. A adição desse conjunto de características no treinamento torna o modelo mais conservador, e um número significativamente menor de contratos, correta ou incorretamente, é classificado como alto risco – apesar de prever menos falsos positivos, o número de verdadeiros positivos também é reduzido.

Na prática, *SAGRES RF Tópicos* seria o modelo em que o menor volume de verdadeiros positivos (6) seria apropriadamente fiscalizado, representando apenas R\$ 6,2 milhões, e o maior volume de falsos negativos seria descartado para um acompanhamento mais próximo (R\$ 190 milhões). Por outro lado, esse também seria o cenário em que a alocação de recursos para o acompanhamento de falsos positivos (16) seria menos necessária, o que representa um montante de R\$ 16 milhões. De modo geral, o resultado da previsão do modelo *SAGRES RF Tópicos* contradiz a opinião de especialistas no que diz respeito à melhoria geral da performance do modelo de risco. Ainda assim, em um cenário com poucos recursos humanos disponíveis para o acompanhamento da execução dos contratos, esse modelo poderia ser útil para a alocação prioritária de recursos. No entanto, deve-se considerar que outros contratos potencialmente problemáticos poderiam não ser identificados.

A segunda característica experimentada, *PIBper capita*, demonstrou possível potencial de melhora no desempenho do modelo. Ao considerar os limites superiores de seus respectivos ICs para a AUC-PR, o modelo *SAGRES RF PIBpc* é 35% superior ao *SAGRES RF*, o que pode representar uma melhora significativa de desempenho nesse cenário. A inclusão dessa característica amplia a capacidade do modelo de identificar a classe de baixo risco sem, necessariamente, impactar significativamente no número de contratos previstos da classe de alto risco. Em termos práticos, ao comparar com o *SAGRES RF*, o modelo *SAGRES RF PIBpc* pode resultar em economicidade, reduzindo os recursos investidos em investigações possivelmente infrutíferas (menor número de falsos positivos, representando R\$ 56 milhões a menos em contratos desse tipo). Além disso, poderia permitir a fiscalização de um montante maior em contratos com sabido risco de rescisão, correspondendo a R\$ 62 milhões a mais, devidamente acompanhados.

O *PIBper capita* pode servir como indicativo do volume de investimento público no próprio município, seja por meio da geração de um maior número de contratos ou contratos de maior porte (para atendimento às demandas populacionais), ou seja, por meio da contratação de pessoal mais qualificado para gestão pública. Em ambos os casos, os mecanismos de controle seriam fortalecidos, tanto pela necessidade de garantia do bom uso do recurso público, quanto pela alta capacitação técnica no planejamento e acompanhamento dos processos. Nesse sentido, a avaliação presente na literatura sobre a contribuição de características relativas ao desenvolvimento econômico do município (ou características análogas, como o capital de compra individual dos cidadãos) em relação ao risco de combate à corrupção é paradoxal [53; 43]. Neste estudo, ainda que o comportamento de interesse possa abranger outros comportamentos além da corrupção (falhas de planejamento ou execução na contratação, por exemplo), a característica econômica do *PIBper capita* demonstrou ser um preditor possivelmente relevante no contexto de contratação pública dos municípios da Paraíba.

Ao considerar o alto custo para reaver recursos desviados de suas finalidades em relação ao próprio valor do contrato [21], e ao considerar as previsões individuais de cada classe e seus respectivos montantes financeiros, é possível que o modelo *SAGRES RF PIBpc* (*pesos obs.*) forneça resultados práticos financeiramente mais favoráveis ao gestor.

Em cenários nos quais ao menos uma das fontes de dados esteja indisponível, observou-se que a ausência das características provenientes do *SAGRES* pode afetar negativamente

o desempenho preditivo dos modelos, enquanto a ausência dos demais conjuntos pode ter impacto meramente marginal. Desse modo, utilizar apenas o conjunto de características latentes Tópicos para modelagem resulta no pior desempenho de classificação de risco.

Comparativamente, a aplicação do modelo *SAGRES RF* resultaria no cenário de maior desperdício de recursos para acompanhamento de contratos possivelmente não problemáticos (falsos positivos), apesar de identificar o segundo maior número de contratos corretamente previstos como alto risco. Já o modelo *SAGRES RF Tópicos* resultaria no cenário de maior montante financeiro em contratos não fiscalizados por erro do classificador (falsos negativos) e menor montante em contratos de alto risco de rescisão devidamente acompanhado (verdadeiros positivos). Por sua vez, o modelo *SAGRES RF (pesos obs.)* resultaria no cenário de maior volume financeiro apropriadamente fiscalizado (verdadeiros positivos) e, também, menor volume financeiro de contratos não acompanhados por erro na classificação (falsos negativos). No entanto, em comparação ao *SAGRES RF PIBpc*, geraria um maior desperdício de recursos no acompanhamento de contratos não irregulares (falsos positivos).

Por fim, a análise do custo-benefício entre o investimento de recursos humanos e financeiros e a identificação de contratos potencialmente problemáticos é uma responsabilidade que cabe à gestão. Para embasar essa análise, o capítulo a seguir examina aspectos gerais e específicos da composição do risco, com base nas características e modelos aqui propostos.

Limitações

Uma limitação deste estudo está relacionada à qualidade da descrição do objeto do contrato, utilizada para gerar a matriz contrato-probabilidades na Seção 7.1. Como esse preenchimento é manual, há risco de descrições incorretas, inconsistentes ou imprecisas, incluindo variações na forma de descrever objetos semelhantes. Nesse cenário, o LDA, que identifica padrões de coocorrência entre palavras para definir a semântica dos tópicos, pode associar contratos similares a tópicos distintos. Como, por exemplo, um objeto que descreva a contratação do aluguel de uma **frota de veículos** para prestação de serviços à secretaria de **saúde ou educação**. Dessa forma, ambiguidades nas descrições podem alterar a distribuição de probabilidade entre os tópicos, impactando a classificação de risco ao associá-los a temas distintos.

Capítulo 8

Explicando o risco de rescisão

No contexto prático da aplicação de modelos de risco de rescisão, a capacidade de explicar e interpretar as previsões geradas é essencial. Isso permite identificar padrões recorrentes, detectar possíveis vieses na classificação dos contratos e diferenciar aqueles que realmente apresentam alto risco. A explicabilidade do modelo possibilita também a descoberta de conhecimento sobre os fatores que influenciam a rescisão contratual, contribuindo para a tomada de decisão mais informada e transparente. Ainda, por se tratar de um tema sensível, pode auxiliar em justificativas para alocação de recursos de fiscalização e gerar maior confiança entre potenciais usuários, incentivando a adoção da ferramenta [39]. Nessa perspectiva, essa análise objetivou inferir características gerais do risco e descrever, em particular, o comportamento de risco dos contratos corretamente classificados pelo modelo.

Selecionou-se o modelo de maior AUC-PR na previsão do conjunto de dados de teste (*Sagres + RF + PIB per capita*) avaliado anteriormente, e, utilizando valores SHAPley, analisou-se a influência (positiva ou negativa) das características nas chances de um contrato ser classificado como de alto risco de rescisão. Assim, a partir disso, foi possível inferir o alinhamento desses resultados com a experiência de especialistas na área, bem como com estudos anteriores.

Como detalhado na Seção 2.6.1, o valor SHAPley de uma característica determina o quanto e de que modo seu valor nessa observação contribuiu para a previsão do modelo em relação ao valor médio previsto. Um valor SHAPley maior que zero indica que o valor de determinada característica eleva o valor de previsão do modelo em relação ao valor médio. Analogamente, um valor negativo indica que o valor dessa característica diminui o valor de

previsão em relação ao valor médio.

A seguir, para interpretação dos gráficos, convencionou-se utilizar a gradação de cores de azul a vermelho, considerando-as cores como canais representativos de valores SHAPley baixos e altos, respectivamente, em relação a cada característica analisada. Em cinza, nos quadrantes inferiores dos gráficos, são exibidos histogramas da distribuição de ocorrência dos dados.

8.1 Influência geral de características

Na Figura 8.1 são exibidas as características ordenadas da maior para a menor influência na previsão de risco. As cinco características de maior contribuição na composição do comportamento de risco de rescisão dos contratos municipais são o PIB *per capita*, a diferença em anos entre a data de abertura da empresa e a data de seu primeiro contrato, o valor do contrato, o montante total ganho pela empresa em contratos até o momento de assinatura daquele contrato, e a razão entre o contrato e o valor recebido.

A seguir, são elencados destaques sobre a influência dos valores de cada característica na modelagem do comportamento de risco de rescisão. É importante destacar que, embora o impacto de cada característica seja analisado de maneira individual, não é esperado que o estado de risco dos contratos seja inteiramente explicado por cada uma delas, podendo haver interação entre subconjuntos de características no processo de aprendizado do modelo.

PIB *per capita*

Com maior influência na previsão de risco, o PIB *per capita* descreve a razão entre montante financeiro arrecadado (por meio de bens e serviços) e a população do município. Visualmente, com base na Figura 8.1, observa-se que, para a maior parte dos contratos, a contribuição dessa característica no comportamento de risco é negativa (valores SHAPley inferiores a zero), com uma tendência de influência não linear em relação aos valores assumidos pela característica. Por outro lado, vê-se que, em contratos nos quais a influência da característica é positiva, o valor do PIB *per capita* é, em geral, alto.

O gráfico de dependências parciais, ilustrado na Figura 8.2, oferece uma visão mais detalhada da relação entre os valores SHAPley e a característica analisada. No eixo x, são

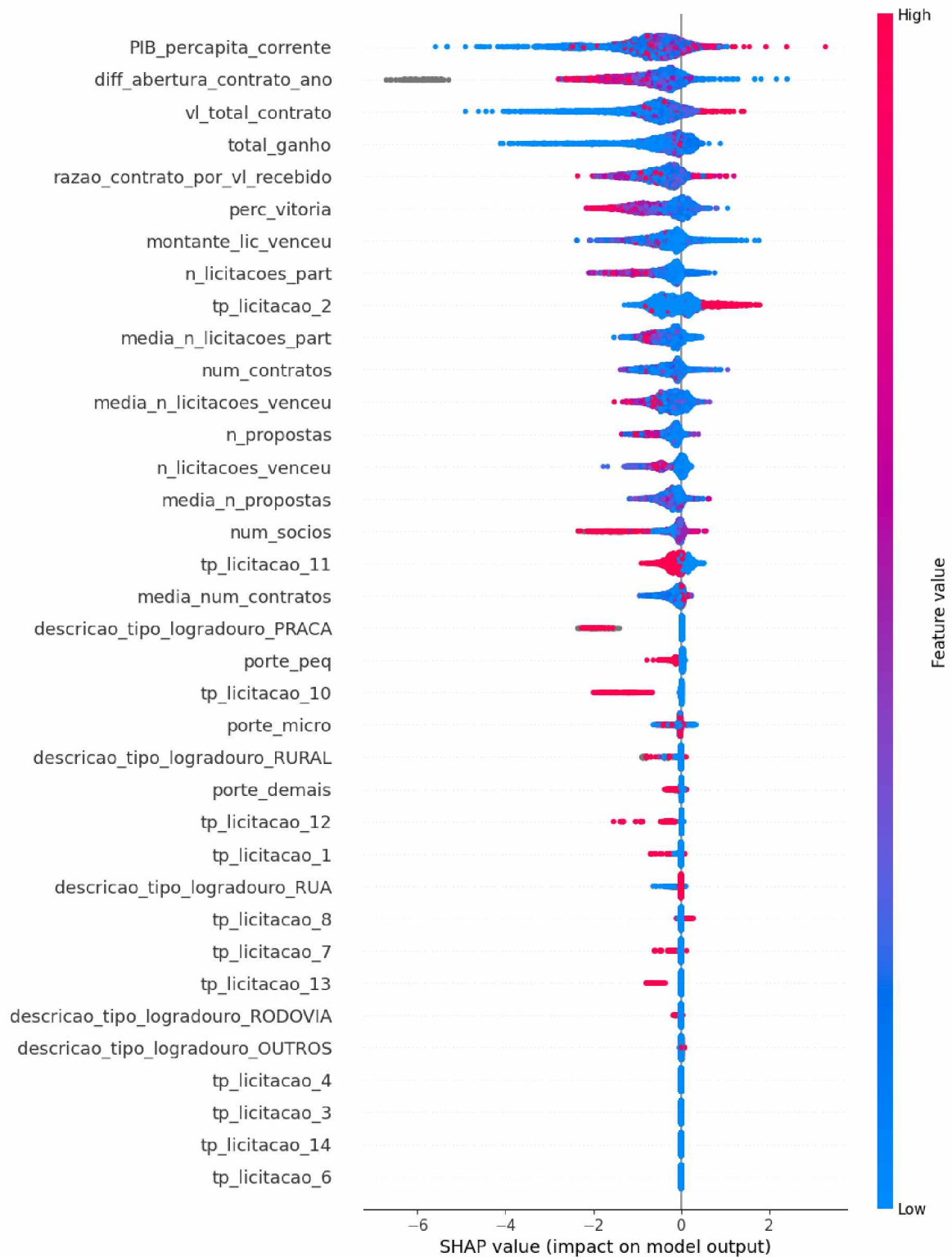


Figura 8.1: Sumário de influência na previsão - Modelo SAGRES + RF + PIB *per capita*

representados os valores assumidos pela característica, enquanto a cor dos pontos reforça essa distribuição. O eixo y exibe os valores SHAPley de cada observação. Uma tendência

ascendente indica que, à medida que a variável cresce, seu impacto no risco aumenta. Da mesma forma, uma tendência descendente sugere que valores menores da característica reduzem sua contribuição para o risco. Essas interpretações lineares são úteis para interpretações locais nas relações não lineares entre valor da característica e contribuição de risco.

Na Figura 8.2, é possível notar quatro grandes grupos. No primeiro grupo, contratos de municípios de PIB *per capita* inferiores a 20 recebem, na grande maioria dos casos, influência negativa desta característica na composição do risco, o que reduz as chances de serem classificados como alto risco de rescisão. Para contratos pertencentes a municípios de PIB *per capita* entre 20 e 35, há uma ambiguidade evidente na influência da característica no comportamento de risco desses contratos: ora exercendo influência positiva, ora negativa. Neste grupo, em específico, nota-se que a medida que o PIB *per capita* se eleva, o volume de contratos cujo risco é influenciado positivamente pela característica é reduzido.

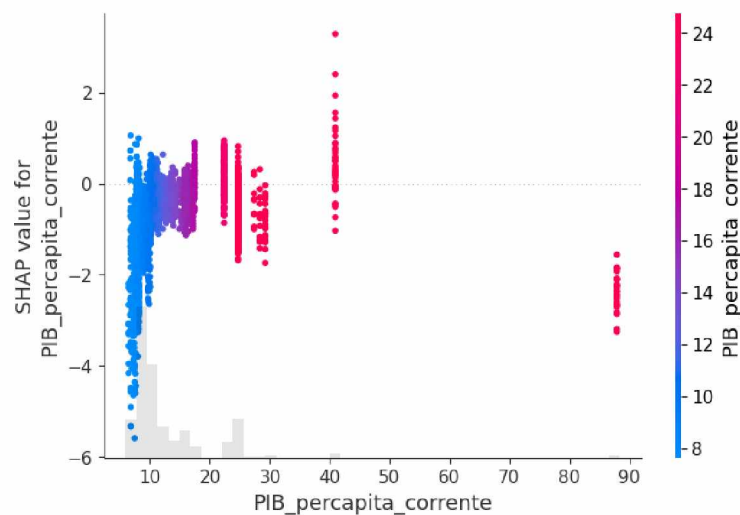


Figura 8.2: Dependências parciais – PIB *per capita*

Um terceiro grupo, composto dos contratos do município de Cabedelo, que possui PIB *per capita* entre 35 e 50, recebem, na grande maioria, influência positiva da característica na composição de seu risco de rescisão. Por outro lado, para a cidade com PIB *per capita* superior a 50, todos os contratos obtêm influência negativa da característica na composição do risco. Especificamente neste grupo, os contratos pertencem a uma mesma cidade (município de Alhandra) e unidade gestora (Prefeitura Municipal), sendo, em alguns casos, originários de uma mesma licitação e, em geral, englobam aquisições de gêneros alimentícios, medicamentos hospitalares e odontológicos, aluguel de veículos ou aquisição de combustível.

Via de regra, é racional supor que a relação de contratos de municípios de baixo PIB *per capita*, com menores chances de rescisão, possa estar associada à também reduzida disponibilidade de recursos humanos e financeiros para o exercício do controle, dado o volume de documentos e os mecanismos de controle reativos empregados. Assim, estando o conjunto de contratos sabidamente irregulares possivelmente subnotificado e sub-representado, esse comportamento será refletido na aprendizagem do modelo acerca do comportamento de risco e, consequentemente, em sua previsão.

Já para contratos de municípios com PIB *per capita* mais elevado, como comportamento geral, ainda quando recebem influência negativa da característica, essa influência tem menor grandeza do que nos casos descritos anteriormente. Logo, é coerente considerar que municípios com PIB *per capita* elevados possuem mais contratos ou contratos de maior porte e, como consequência, dispõe de melhores condições e recursos para exercer controle sobre esses instrumentos – aumentando a chance de que um maior número de irregularidades seja identificado e registrado. Nesse caso, o viés inerente ao funcionamento do próprio mecanismo de controle favorece a construção de um conjunto de rótulos mais representativo do comportamento de risco contratual real, reforçando a ideia de que a capacidade de previsão do risco de rescisão está diretamente ligada à capacidade de controle do município.

Diferença entre o ano de abertura da empresa e o ano do contrato corrente

A distribuição dos valores SHAPley na Figura 8.3 indica um impacto mais linear dessa característica na definição do risco de rescisão dos contratos, especialmente quando comparada ao comportamento de outras características avaliadas. Nesse sentido, considera-se empresas jovens aquelas que possuem data de registro próxima à data de assinatura do seu primeiro contrato, enquanto empresas maduras possuem uma janela temporal maior entre essas duas datas.

Apesar de grande parte dos contratos receber influência negativa da característica, vê-se que contratos de empresas com mais de 30 anos de atuação, por exemplo, estão majoritariamente associados a uma influência negativa expressiva e influência positiva marginal. Já a influência da idade em contratos de empresas abertas há menos de 30 anos em relação à data do contrato é menos constante, podendo ser equilibrada e expressivamente positiva e negativa para empresas até cinco anos, e majoritariamente negativa a partir dessa idade.

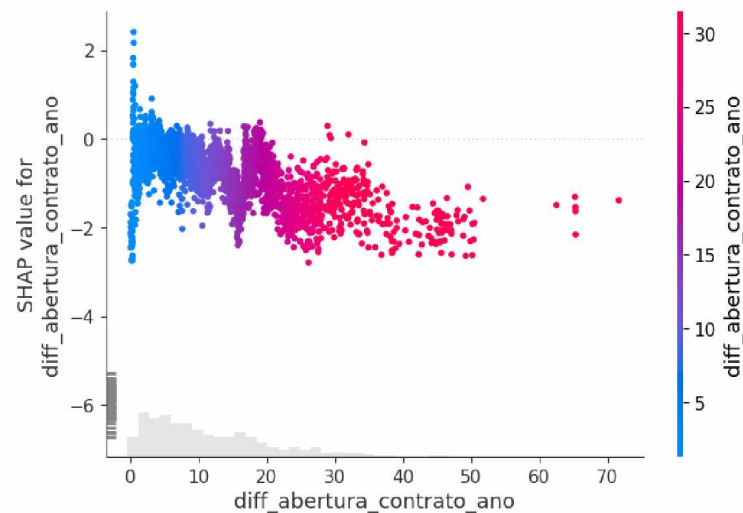


Figura 8.3: Dependências parciais – Diferença entre o ano de abertura da empresa e o ano do contrato corrente

Desse modo, o impacto positivo da característica nos contratos de empresas jovens reforça a percepção comum entre promotores de que empresas com menos experiência podem ter menor capacidade de cumprimento dos seus deveres contratuais, seja por planejamento insuficiente, pouca experiência com processos públicos ou mesmo competências técnicas reduzidas. Ainda, ratifica a percepção de que, em alguns casos, a criação de empresas pode ser movida por finalidades particulares ou interesses suspeitosos, podendo, porventura, resultar na rescisão futura de seus contratos. Paralelamente, o impacto majoritariamente negativo da característica em empresas maduras também fortalece o pressuposto de que empresas com maior experiência no mercado representem, em geral, um risco menor à AP.

Valor do contrato em reais

É importante lembrar que, porque o conjunto de contratos é diverso em temas, há ocorrência de valores variados para o valor do contrato, como nota-se no histograma auxiliar presente nos quadrantes inferiores da Figura 8.4. A distribuição dos valores SHAPley em relação ao valor do contrato é dicotômico: tanto o conjunto de contratos de alto quanto de baixo valor podem obter contribuições positivas ou negativas da característica no risco de rescisão.

Na Figura 8.5, após aplicação da transformação logarítmica de base 10 no eixo do valor

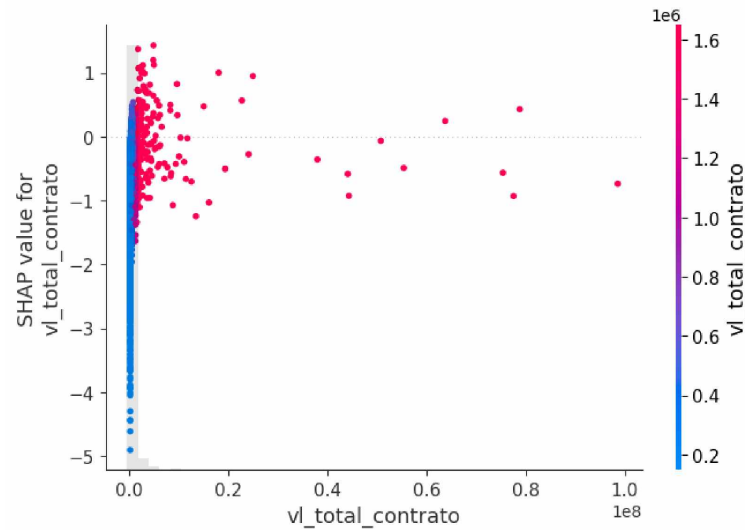


Figura 8.4: Dependências parciais – Valor do contrato em reais (R\$)

do contrato, fica evidente o crescimento do impacto da característica no risco de rescisão dos contratos. Nesse viés, valores baixos de contratos tendem a contribuir negativamente na composição do risco de rescisão desses contratos, e essa contribuição aumenta à medida que o valor do contrato cresce. Para contratos de alto valor, há um maior equilíbrio entre o número de contribuições positivas e negativas na composição do risco, indicando que outras características podem ter influência na composição do risco para essa parcela de contratos.

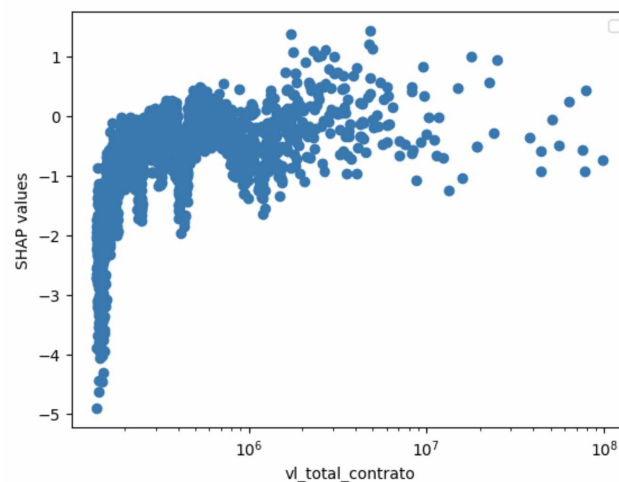


Figura 8.5: Valor do contrato em reais (log10) em relação aos valores SHAPley

Assim, empresas cujos contratos recebem contribuições positivas no risco para essa característica são, na maior parte dos casos, construtoras e empresas de engenharia, postos de

combustíveis, laboratórios e fornecedores de material médico-hospitalar. Em geral, esses ramos de atividade possuem contratos de valores vultuosos, nos quais irregularidades são identificadas com maior frequência e, muitas vezes, percebidas de modo mais evidente pela própria população — como no caso de obras inacabadas ou não iniciadas e atendimento de saúde precário, por exemplo.

Um maior impacto dessa característica em contratos de maior valor, em detrimento do restante do conjunto de contratos, pode também ser reflexo de um mecanismo de controle que investiga principalmente contratos de maior volume monetário, nos quais o investimento de recursos humanos e financeiros teria um bom custo-benefício para a AP, especialmente em casos em que seja possível mitigar prováveis irregularidades. Ainda, considerando que o impacto de cada característica não é um produto isolado, mas sim um reflexo da combinação entre esta e outras características, presume-se que a dualidade de interpretação da influência pode ocorrer devido à interação desta com as demais características.

Razão entre o valor do contrato e o total recebido

Essa característica denota a frequência de atividade da empresa para com a AP até o momento de assinatura do contrato, diferenciando fornecedores frequentes de esporádicos.

Uma razão que se aproxima, ou é igual a 1 (um), representa que o valor do contrato em questão é próximo ou estabelece todo o valor arrecadado pela empresa, caracterizando-a como fornecedora esporádica. Por outro lado, quanto mais próxima de zero a razão estiver, menor a representatividade desse contrato no total recebido pela empresa, seja por seu valor pequeno comparado aos demais, ou pela grande quantidade de contratos firmados (alta frequência de fornecimento). O histograma nos quadrantes inferiores da Figura 8.6 exibe a distribuição da quantidade de contratos por faixa de razão calculada, demonstrando que o conjunto de dados é composto majoritariamente por contratos de fornecedores frequentes ou contratos de baixo valor (baixo valor da característica).

Nota-se que contratos de fornecedores frequentes (ou contratos de baixo valor) recebem, predominantemente, contribuições negativas da característica na composição do risco, isto é, têm suas chances de serem classificados como alto risco de rescisão diminuídas. À medida que o valor da razão se eleva, esse comportamento se intensifica e se mantém até certo ponto. Porém, entre 0,15 e 0,25, há um grupo de contratos que recebe exclusivamente influência

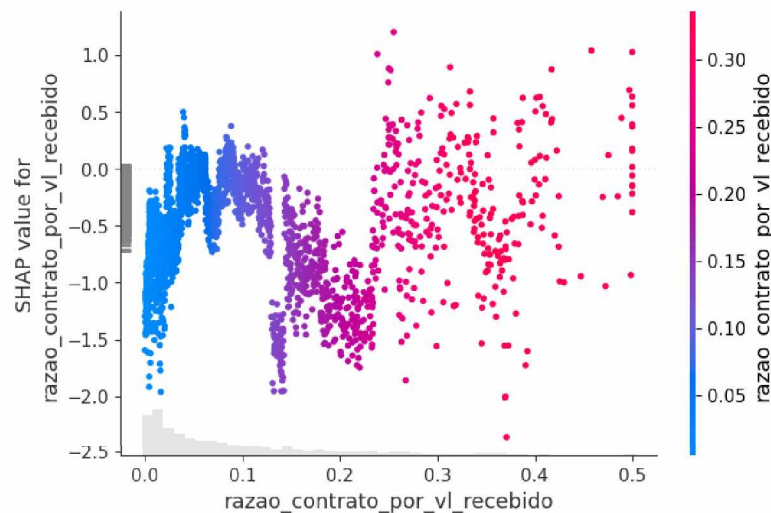


Figura 8.6: Dependências parciais – Razão entre o valor do contrato e valor total recebido

negativa da característica. Contratos de fornecedores esporádicos (ou contratos de maior valor) ocorrem em menor frequência e obtêm influência positiva ou negativa na composição do seu risco de rescisão, de forma ambivalente.

Por meio do gráfico de dependências parciais, também é possível relacionar o comportamento de duas características à contribuição de uma delas na composição do risco de rescisão. Na Figura 8.7, a *razão entre o valor do contrato e total recebido* e o *número de propostas feitas pela empresa em licitações* foram associados com a intenção de observar se proponentes frequentes são, também, fornecedores frequentes.

O canal da cor identifica, na Figura 8.7, os contratos de proponentes frequentes (pontos vermelhos) e de proponentes não frequentes (pontos azuis). É possível observar, com poucas exceções, que proponentes frequentes são também fornecedores frequentes e, em sua maioria, recebem influência negativa no risco proveniente da segunda característica – oferecendo menos chances de rescisão contratual. Diante disso, contratos de pequeno valor ou de fornecedores frequentes e que não propõem com frequência tendem a estar relacionados com maiores chances de rescisão contratual.

A baixa frequência de fornecimento (ou um alto valor do contrato em relação aos demais dessa empresa) por um fornecedor com um grande número de propostas (pontos vermelhos à direita do gráfico) não apresenta uma relação direta com o risco de rescisão. Apesar disso, ainda é um comportamento valoroso, uma vez que pode estar relacionado a outra prática comum: a criação de empresas de fachada que participam da licitação e fazem propostas sem

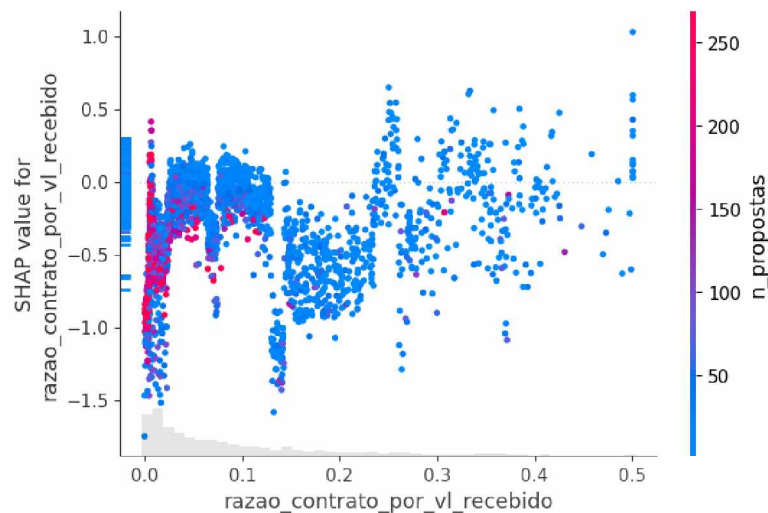


Figura 8.7: Dependências parciais – Relação entre características (*razão entre o valor do contrato e valor total recebido* vs. *número de propostas feitas pela empresa em licitações*)

intenção de vencer, com a intenção de comprometer o caráter concorrencial do processo.

Outras características

À medida que o total (monetário) ganho pela empresa em contratos aumenta, sua contribuição no comportamento de risco também tende a aumentar. Essa característica contribui positivamente para o comportamento de risco de cerca de 25% de todos os contratos. Dentre os 25%, apenas 4% pertencem a empresas com faturamento de até cinco milhões de reais no período compreendido. Empresas de baixo faturamento podem representar pequenos produtores, micro e pequenos fornecedores, agricultores familiares, por exemplo. Ademais, contratos provenientes da modalidade de licitação dois (tomada de preços) obtêm influência positiva na composição do risco em 79% dos casos.

8.2 Características relevantes para contratos verdadeiros positivos

Na seção anterior, foi analisado como as mudanças nos valores e distribuições de cada característica impactam a previsão de risco de rescisão fornecida pelo modelo para todo o conjunto de contratos disponível para a modelagem, o que fornece um panorama tanto da influência

das características no comportamento de risco quanto da dicotomia dessa influência na composição do risco de rescisão. Para aprofundar a compreensão acerca do comportamento de risco de rescisão em contratos através de classificações específicas, foram selecionados apenas os contratos verdadeiros positivos (contratos de alto risco classificados corretamente pelo modelo). Com isso, é possível analisar a influência das características na composição do risco para cada contrato verdadeiro positivo de maneira individual.

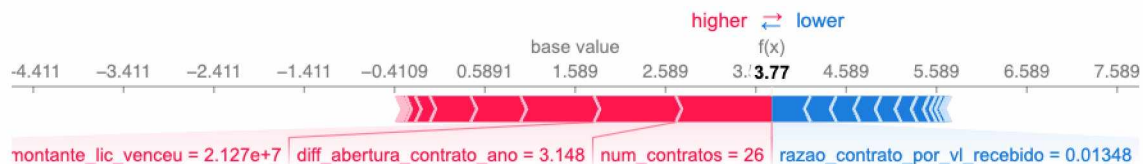


Figura 8.8: Gráfico de força individual – Exemplo de contrato verdadeiro positivo

O valor base (*base value*) é a média das previsões do modelo no conjunto de dados de treinamento (isto é, o valor esperado da previsão). A contribuição de cada característica é computada com base no aumento ou diminuição do valor base, enquanto o valor previsto ($f(x)$) é calculado a partir da soma do valor base às contribuições de cada característica. Assim, cada característica é representada, na Figura 8.8 (gráfico de força), por um bloco.

O tamanho de cada bloco indica a magnitude do impacto da característica, ao passo que setas em vermelho denotam contribuição positiva e setas em azul indicam contribuição negativa para a classificação de risco de rescisão do contrato. Características em vermelho aumentam as chances de que o contrato seja classificado como alto risco de rescisão, enquanto características em azul reduzem as chances de que o contrato seja classificado como alto risco de rescisão. Além disso, características mais relevantes estão posicionadas ao centro do gráfico, seguidas, em ordem de relevância, pelas demais, ordenadamente, à esquerda e à direita. A partir dessa visualização, é possível compreender como a "força" (impacto) de cada característica afeta o valor previsto pelo modelo para cada contrato individualmente. Para cada contrato, é possível identificar características com influência comum, e, a seguir, são expostas e analisadas as que possuem contribuições expressivas no risco.

O PIB *per capita*, característica de destaque com maior influência positiva no risco de rescisão, como visto em 8.1, possui impacto em 14 dos 16 contratos de alto risco de rescisão (87%) e é a característica mais influente na classificação de risco de seis deles. Enquanto isso, a diferença em anos entre a abertura da empresa e seu primeiro contrato representa a segunda

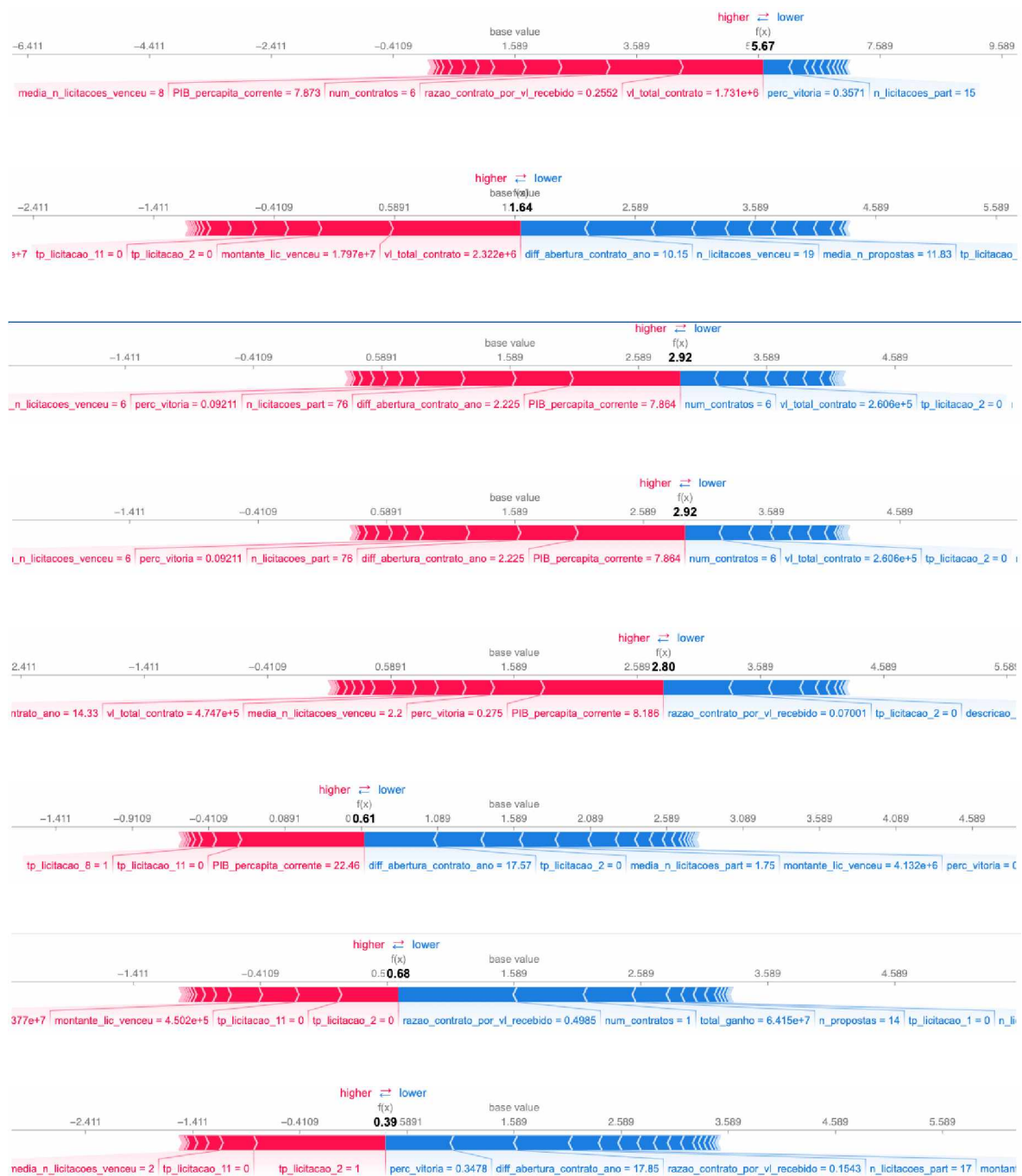


Figura 8.9: Gráficos de força – Contratos verdadeiros positivos (parte D)

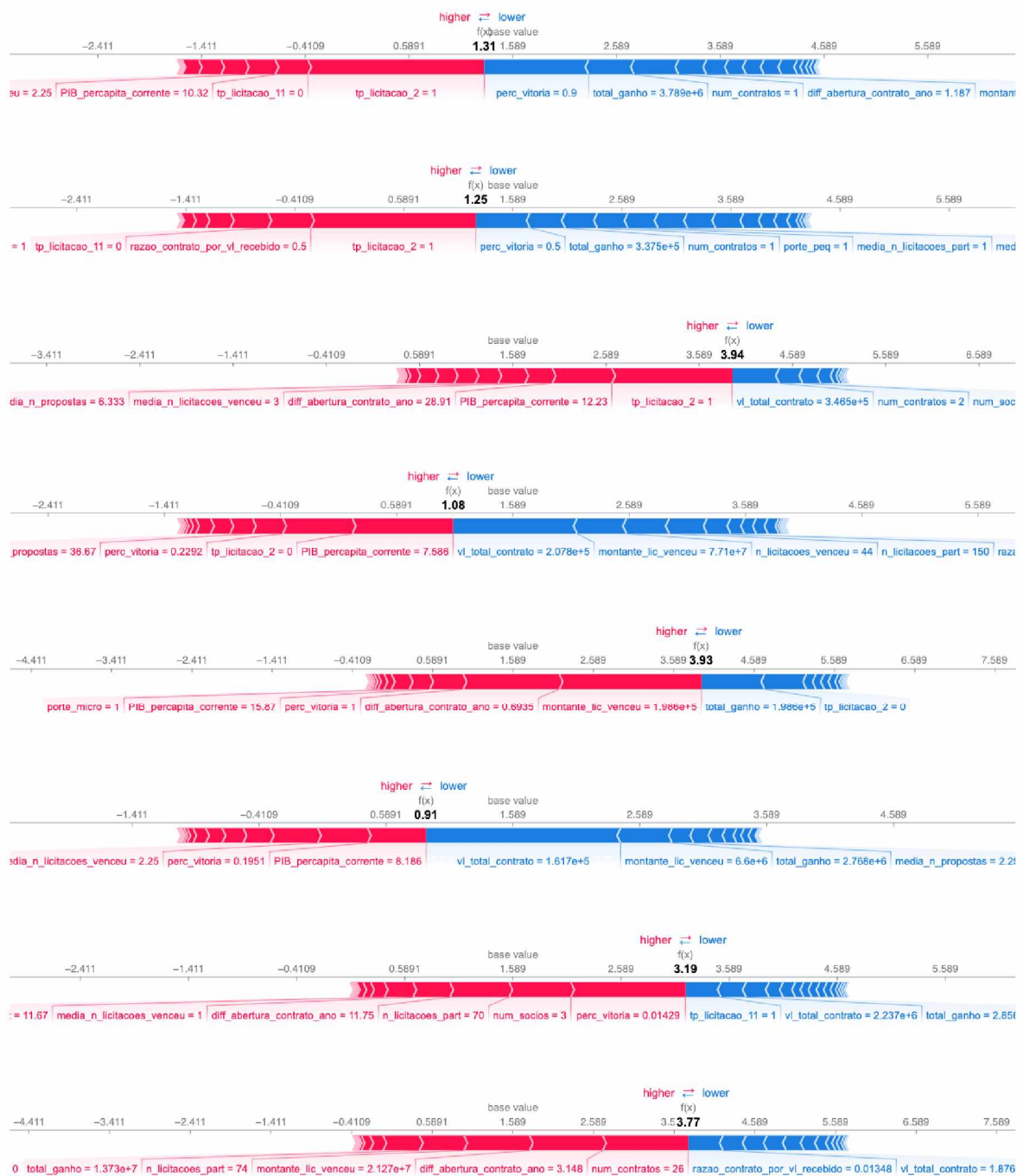


Figura 8.10: Gráficos de força – Contratos verdadeiros positivos (parte II)

maior influência positiva e maior influência absoluta, com influência positiva em 10 dos 16 contratos verdadeiros positivos, mas sem demonstrar contribuição linear na classificação dessa amostra.

Considerando contratos de empresas com tempo de atuação de zero a cinco anos, cinco dos oito contratos recebem influência positiva em seu risco de rescisão. Isso pode significar que, em uma eventualidade, empresas jovens em atuação podem ter dificuldades em cumprir obrigações contratuais, o que representa um risco ao fornecimento do bem ou produto adquirido, por exemplo. Para contratos de empresas entre 10 e 20 anos, dois dos seis contratos são positivamente influenciados pela característica. Já para contratos de empresas acima de 20 anos, há influência positiva da característica na classificação de risco de rescisão em todos os dois contratos. É esperado que empresas experientes, com muitos anos de atuação, representem um risco menor à AP, o que contradiz a influência positiva da característica no risco dos contratos de empresas com mais de 20 anos.

Para contratos até R\$ 300.000,00, em que a característica valor total do contrato imprime efeito positivo (seis contratos), quanto maior o valor do contrato, mais significativa aparenta ser a influência da característica nas suas chances de rescisão. Por outro lado, para contratos entre R\$ 300.000,00 e R\$ 500.000,00, em oito dos 10 contratos a característica tem efeito negativo no risco – quanto menor o valor do contrato, maior o efeito negativo da característica no risco de rescisão, no geral.

O total monetário ganho em contratos tem um efeito negativo na composição do risco de oito contratos, enquanto o número de contratos de cada empresa impacta negativamente 10 dos 16 contratos com alto risco de rescisão. Assim, ambas as características não imprimem efeito constante no comportamento de risco, mesmo para faixas de valores semelhantes.

Apenas cinco, dentre os verdadeiros positivos, recebem influência positiva da característica "razão entre o valor do contrato e o valor total recebido" na composição do risco de rescisão. Desses, apenas dois são de empresas cujos contratos em questão representam 25% ou mais de toda sua arrecadação em contratos (fornecedores esporádicos ou contratos de alto valor). Por outro lado, as contribuições negativas da característica aos oito contratos de empresas cujo contrato em questão representa 20% ou menos do valor recebido (fornecedores frequentes ou contratos de baixo valor) são mais consistentes, o que pode ser explicado pela expertise do fornecedor com os procedimentos da AP ou menos interesse, em geral, nesse

tipo de contrato.

Na característica "percentual de vitórias", um valor alto pode representar empresas que participam de muitas licitações e vencem muitas delas, ou empresas que participam pouco, mas ganham em proporção similar. De modo geral, a influência dessa característica tende a diminuir à medida que o percentual de vitórias aumenta. Contratos de empresas que possuem entre 1 e 30% de vitórias tendem a receber contribuições positivas da característica no cálculo do risco de rescisão.

O montante vencido em licitações imprime efeito positivo no risco de oito contratos dentre os verdadeiros positivos. Ainda que essa seja uma característica relevante na classificação de risco de rescisão de quatro dos oito contratos classificados corretamente pelo modelo, não há relação aparente entre o valor da característica e sua contribuição para o risco, mesmo em faixas de valores semelhantes.

Da característica "número de contratos" firmados por cada empresa, 10 contratos recebem influência negativa. Seis dentre os oito contratos de empresas que possuem entre um e cinco contratos obtêm influência negativa da característica na composição do risco. Na faixa de empresas que possuem entre 6 e 10 contratos, a influência da característica é ambivalente; dois dos quatro contratos obtêm influência negativa. Para empresas com mais de 10 contratos, o impacto da característica é positivo e relevante. Apesar de não haver uma contribuição linear, ao considerar as faixas de valores da característica, o impacto positivo no cálculo do risco tende a aumentar à medida que o número de contratos também cresce.

Para a modalidade de licitação dois (tomada de preços), todos os quatro contratos que advêm dessa modalidade recebem efeito positivo na classificação de risco de rescisão. Além disso, quatro dentre os contratos firmados a partir de outras modalidades de licitação recebem contribuições positivas da característica na composição do risco de rescisão, aparente indicativo da influência de outras características na composição do risco.

Para sete dos oito contratos provenientes da modalidade de licitação 11 (pregão presencial), há contribuição negativa da característica no risco de rescisão. Na prática, essa modalidade de licitação é historicamente conhecida por especialistas da área por casos de intimidação e fraude, uma vez que os licitantes interessados devem apresentar-se de maneira presencial para registrar seus lances. A partir disso, é contraintuitivo que essa característica contribua negativamente em tantos casos na composição do risco de rescisão, enquanto os

demais contratos não associados a ela são, majoritariamente, positivamente influenciados quanto ao risco de rescisão.

A concorrência, modalidade de licitação um, frequentemente utilizada para contratações de maior valor (em geral, obras e serviços), tem influência negativa em todos os 16 contratos corretamente classificados como alto risco de rescisão, tendo, nos dois contratos provenientes dessa modalidade, um impacto negativo ainda maior que nos demais.

Assim, as características da modalidade de licitação 3 (convite), 4 (concurso), 6 (dispensa por valor), 7 (dispensa por outros motivos), 10 (pregão eletrônico), 12 (adesão a registro de preço), 13 (chamada pública), 14 (regime diferenciado de contratação) e tipo do logradouro, como praça ou rodovia, não possuem ocorrência no conjunto de dados de teste, tendo contribuições marginais, positivas ou negativas, no comportamento de risco de rescisão dos contratos. Um conjunto de contratos mais abrangente é necessário para determinar a influência dessas características no risco de rescisão.

8.3 Discussão

A compreensão do impacto de cada característica na composição do risco é fundamental para compreender o comportamento preditivo do modelo e assegurar que o comportamento de interesse dos contratos está sendo apropriadamente construído a partir das características escolhidas como predictoras. Por tratar-se de um tema delicado, esse tipo de entendimento também fornece segurança ao público de interesse, em especial aos órgãos de controle e à sociedade civil, ao adotar uma ferramenta que pode gerar impacto em sua dinâmica de controle, atuação interna e externa e, ainda, em casos extremos, replicar comportamentos investigativos históricos enviesados.

Modelar o comportamento de risco de rescisão é uma tarefa complexa e, como observado, não há característica única dentre as utilizadas neste trabalho que seja um descritor constante ou exclusivo do comportamento de risco de rescisão. Entretanto, é possível identificar características frequentemente relevantes e padrões de comportamento de influência na composição do risco a partir da modelagem, sendo ainda possível comparar tendências em um cenário que compreende todo o conjunto de dados com um mais específico, que abarca o subconjunto de contratos corretamente classificados como alto risco pelo modelo. À vista

disso, algumas características foram destaque na análise da composição do risco de rescisão dos contratos de municípios paraibanos.

Nessa perspectiva, cidades com menor PIB *per capita*, via de regra, dispõem de menos recursos financeiros e menor capacidade administrativa para garantir a qualidade dos serviços e um controle efetivo e compreensivo para o volume de contratos firmados. Por conseguinte, os mecanismos de transparência e o engajamento da sociedade civil, fortes aliados ao controle reativo de órgãos públicos, ativado a partir de potenciais irregularidades ou denúncias, também podem ser fragilizados pela falta de investimentos, pelo desconhecimento dos processos públicos ou mesmo pela desmotivação dos cidadãos pela percepção de desonestidade dos governantes [42]. Fatores como esses, em cidades com menor arrecadação, podem aumentar as chances de que, por descuido, omissão ou até mesmo conduta irregular intencional por parte da empresa contratada ou da AP, as respectivas controladorias ou cortes de contas sejam convocadas a intervir.

Por outro lado, ainda que com maior número de contratos e possivelmente contratos mais vultuosos, cidades com maior PIB *per capita* tendem a contar com estruturas organizacionais e de governança mais robustas, investindo em capacitação técnica de servidores e em ferramentas para auxiliar na gestão e fiscalização de seus processos [50, Cap. 2]. Contratos de valores altos também podem gerar maior interesse em empresas licitantes, com interesses escusos ou não, o que pode intensificar o crivo no monitoramento da execução desses instrumentos. Assim, contratos de cidades com PIB *per capita* e, consequentemente, qualidade de governança e participação popular maiores, podem contar com sistemas, ferramentas ou pessoal próprio para identificação e correção de falhas ou irregularidades no planejamento ou percurso de sua execução, o que pode reduzir as chances de que esses contratos caiam em escrutínio futuro pelos órgãos de controle e venham a ser rescindidos.

Na modelagem, contratos de cidades com menor arrecadação sofrem influência predominantemente negativa dessa característica em relação ao risco de rescisão. Já para contratos de cidades com maior arrecadação, o impacto da característica demonstra um comportamento mais equilibrado e ambivalente, o que indica uma interação mais complexa e possivelmente relacionada a outros fatores associados ao risco. Em contrapartida, na análise dos contratos corretamente classificados como alto risco de rescisão pelo modelo, os comportamentos descritos ficam evidentes para contratos firmados em cidades de PIB *per capita* baixo, onde a

característica imprime influência positiva em nove dos 10 contratos dessas cidades.

Para os contratos de cidades com PIB *per capita* moderado ou alto, a influência da característica é majoritariamente positiva (em cinco de seis contratos). Em contraste com a expectativa, é possível que essa influência esteja relacionada à existência de órgãos de controle com estruturas mais robustas, apesar da presença de processos governamentais locais mais desenvolvidos, possibilitando a identificação de um maior número de contratos cujo planejamento ou execução envolveram falhas ou irregularidades.

O efeito da experiência das empresas, tanto na qualidade do serviço prestado à AP, quanto nas chances de que estas estejam sujeitas a penalidades graves, foi discutido em estudos anteriores [6; 46]. Destaca-se que empresas mais experientes tendem a possuir maior maturidade organizacional e maior conhecimento acerca dos processos de gestão e das necessidades do setor público, atingindo os objetivos contratuais de maneira satisfatória e, em geral, com qualidade superior. Além disso, também possuem a capacidade de apresentar propostas mais alinhadas com as especificações demandadas [11; 46], resguardando interesses próprios, mas também alinhados com os interesses da AP.

Nesse contexto, em todo o conjunto de dados, a diferença entre o ano de abertura da empresa e o ano em que o contrato está sendo firmado (idade da empresa) denota a maturidade e a experiência da empresa, o que impacta de maneira negativa o risco de contratos de empresas mais experientes. Esse comportamento está alinhado, sob uma perspectiva de risco contratual e modelagem empírica, com as conclusões dos estudos citados previamente. Em geral, quanto maior for a idade da empresa, menor será a influência da característica na composição do risco de seus contratos.

Já a ambiguidade da influência da idade de empresas mais jovens na composição do risco de rescisão de seus contratos pode refletir a heterogeneidade na composição do risco de rescisão desses contratos, uma vez que outros fatores podem ser determinantes para seu sucesso ou insucesso, e a não falha do modelo em identificar essa diversidade de comportamentos. A própria capacidade de gestão e experiência dos agentes públicos envolvidos no acompanhamento da execução dos contratos, por exemplo, podem ser fatores relevantes para a conclusão exitosa do processo de prestação de serviço [11]. Ademais, limitações competitivas e financeiras, a inexperiência na gestão contratual ou mesmo a dificuldade em lidar com idiossincrasias da contratação pública podem contribuir para o fracasso na contratação

de empresas de menor idade.

Nos contratos verdadeiros positivos, essa característica contribui para o risco de 10 dos 16 contratos corretamente classificados pelo modelo, independente do grau de maturidade da empresa responsável. Entretanto, fica evidente que, para empresas mais jovens, seu impacto no risco de rescisão dos contratos pode ser mais expressivo.

Nesse estudo, utiliza-se a quantidade de anos de atuação da empresa de modo abrangente, sem diferenciação entre atuação no setor público, privado ou mesmo o segmento de atuação da empresa (indústria, comércio etc.). A partir do comportamento dessa característica e dos estudos citados, é possível que a estratificação da experiência nos estratos citados, por exemplo, com a intenção de fornecer ainda mais consistência ao constructo, possa criar preditores ainda mais influentes no risco.

A frequência de contratação de uma empresa também é um indicativo de sua experiência e capacidade na prestação de serviços ao governo e à população. É razoável que empresas frequentemente contratadas pela AP estejam familiarizadas com processos de gestão, normas e regulamentações. A característica que define a razão entre o valor do contrato e o valor ganho pela empresa em outros contratos com a AP demonstra a importância que um contrato específico possui dentro o montante total recebido pelo engajamento em outros contratos públicos. Ainda, essa característica também pode denotar a proporção do valor de um determinado contrato em relação aos demais firmados pela mesma empresa.

Ao considerar o conjunto total de contratos, a frequência de fornecimento (ou tamanho do contrato) demonstra alinhamento com o comportamento da característica descrita anteriormente e de achados dos estudos mencionados, em especial nos contratos de fornecedores mais frequentemente contratados. Não obstante, ao receberem, de modo mais equilibrado, influências positivas e negativas na composição do risco de rescisão, o comportamento de risco de contratos de empresas engajadas em menos contratos também é condizente com o de empresas jovens. Esses comportamentos são corroborados, inclusive, nos contratos cujo risco foi corretamente classificado pelo modelo para ambas as direções de influência no risco.

Em se tratando do valor do contrato, a possibilidade de altos ganhos financeiros pode ser um atrativo para a participação de fornecedores oportunistas nos processos de licitação e contratação [17]. Empresas que assumem contratos de maior valor possuem maiores chances de receber sanções, em especial em casos em que há colusão entre empresas. Já contratos

de menor valor receberam pouca atenção em debates de estudos anteriores, não havendo resultados conclusivos acerca de sua relação com casos de corrupção ou possíveis sanções às empresas relacionadas.

Na explicação da modelagem de risco para o conjunto de contratos municipais, quanto menor for o valor do contrato, mais razoavelmente expressiva é a contribuição negativa da característica na composição do risco. À medida que o valor do contrato aumenta, de modo geral, aumentam também os valores SHAPley (o que pode representar aumento da contribuição positiva ou diminuição da contribuição negativa). Esse impacto corrobora a literatura teórica quanto ao risco de contratos de alto valor e revela o comportamento para contratos de baixo valor.

Ainda, é possível que, uma vez que os mecanismos de controle empregados por governos e órgãos de controle são comumente reativos, essa associação pode ser justificada pelo viés de seleção no monitoramento de contratos. Aqueles de maior valor seriam acompanhados com maior cautela e proximidade, na medida em que o montante de recursos passível de perda seria ainda mais significativo, propiciando a identificação e notificação de um maior número de contratos irregulares. Um paralelo a essa conjectura também pode ser aplicado aos contratos menores que, por sua vez, teriam suas irregularidades identificadas em menor proporção. Além disso, é possível destacar que, a depender da modalidade de licitação originária do contrato, empresas podem atuar em conluio no processo de licitação para elevar o preço do objeto ou serviço prestado, levando a maiores ganhos para o licitante vencedor. As demais características não demonstraram impacto expressivo no comportamento de risco.

8.4 Limitações

As limitações desta análise estão associadas à subnotificação e sub-representação de casos positivos nos dados utilizados como rótulos de risco de rescisão, o que pode resultar em uma compreensão parcial do comportamento de risco. A análise dos casos corretamente classificados pelo modelo fica, então, restrita a contratos previamente investigados, cujo comportamento de risco foi previamente identificado pela gestão ou órgãos de controle. Além disso, é importante observar que o instrumento do contrato pode ser substituído, em certos casos, por documentos similares, como notas de empenho, autorizações de compra, cartas-contrato ou

ordens de execução de serviço [8, Art. 62]. Para viabilizar a replicação e comparação dos resultados e das análises em outros contextos, além de simplificar a geração de características, apenas contratos foram considerados nesse estudo.

Capítulo 9

O risco de rescisão na prática

No capítulo anterior foi explorada a influência de características relevantes para a composição do risco de rescisão dos contratos, além de detalhar como esse risco foi efetivamente constituído para cada um dos contratos corretamente classificados pelo modelo de AM.

Este capítulo tem o propósito complementar de avaliar o alinhamento do comportamento de aprendizado do modelo de classificação de risco com as justificativas presentes nos pareceres técnicos das rescisões (apresentados pelos auditores em contas públicas) ou com tipologias de risco. À vista disso, objetiva-se validar o alinhamento entre a semântica das características preditoras do risco e os acontecimentos reconhecidamente irregulares ou suspeitos pelos especialistas. Por fim, valer-se do histórico de execução de cada contrato-exemplo para identificar possíveis novos preditores de risco que possam ser experimentados em trabalhos futuros.

A partir da tramitação de documentos e processos relacionados aos contratos a partir do portal do Tramita, do TCE-PB, foram explorados detalhes dos documentos contratuais em si, suas execuções reais e as especificidades de seus encerramentos. Para isso, quatro contratos com maiores chances de rescisão atribuídas pelo modelo foram selecionados. Ainda, um contrato falso positivo e um falso negativo também foram incluídos, de modo a explorar falhas e comportamentos inesperados na estimação do risco de rescisão.

Uma vez que o objetivo desta análise não é expor as partes envolvidas nos contratos, mas sim avaliar o comportamento preditivo do modelo em relação à realidade, foram omitidos identificadores dos contratos, documentos auxiliares, contratantes e contratados.

9.1 Contrato I - Construção de unidades básicas de saúde

O primeiro contrato analisado obteve 99,6% de chances de rescisão pelo modelo. O contrato teve início em 09/04/2014 com vigência inicial até 05/12/2014, e abrangia a *"Contratação de Empresa Especializada para Construção de 04 (quatro) Unidades Básicas de Saúde/UBS Porte I no Município (...)"* no valor total de R\$ 1.730.622,00 – que representa cerca de 25% do total de faturamento da empresa em contratos com a AP. A modalidade de licitação originária do contrato foi a concorrência. No decorrer da execução do contrato, foram executados dois termos aditivos de prazo e um aditivo de supressão no valor do contrato (ao que consta, por entendimento do setor técnico de engenharia da prefeitura contratante), e empenhados mais de R\$ 940 mil reais para realização de pagamentos. A empresa contratada registrou o início das suas atividades na RF em novembro de 2009, consta em cadastro como de pequeno porte e estava localizada em um município a 97 quilômetros da execução da obra.

Alguns dos fatos elencados podem levantar dúvidas acerca da capacidade técnica da empresa e viabilidade da execução do contrato, como: um contrato de grande valor (quase três vezes superior ao terceiro quartil dos valores dos contratos presentes no conjunto de testes) na área de construção ou engenharia (área cujos contratos estão, tipicamente, mais suscetíveis a problemas ou falhas em suas execuções) sendo assumido por uma empresa de pequeno porte, relativamente jovem (aberta apenas quatro anos antes da assinatura do contrato) e com poucos contratos anteriormente firmados com a AP. Segundo a justificativa da rescisão contratual publicada no Tramita, *"(...) Pelo fato da empresa contratada ter paralisado sem justa causa e sem comunicação prévia a execução da obra, levando a administração a comprovar a impossibilidade da conclusão da obra, gerando o descumprimento das cláusulas contratuais, notadamente a que se refere ao prazo de execução (...)"*.

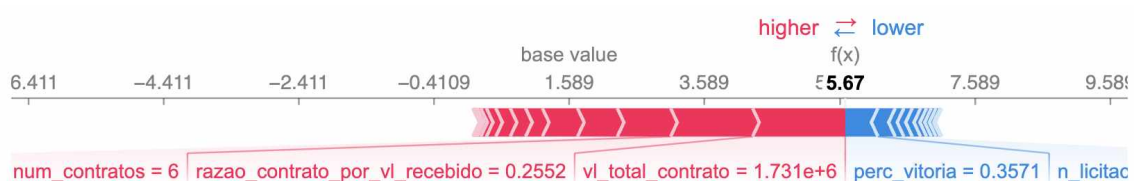


Figura 9.1: Composição do risco de rescisão – Contrato I

Sendo o valor do contrato um indicativo de complexidade ou de dimensão de seu objeto, para empresas de pequeno porte, é comum que contratos de alto valor sejam assumidos por

empresas com um reduzido número de contratos anteriores, sendo possível que isso aconteça devido à necessidade do comprometimento de recursos e de capacidade técnica disponíveis com um único contrato por vez. O número de contratos anteriores da empresa foi inferior à média de contratos de empresas desse porte (13).

O modelo foi capaz, através do conjunto de características utilizado, de capturar alguns desses comportamentos, como observado na Figura 9.1. As três características mais relevantes na composição do risco positivo de rescisão podem, nesse caso, denotar a complexidade do contrato em e si, a dimensão desse contrato em relação aos demais já executados e a inexperience histórica da empresa em contratos anteriores, firmados com a AP, respectivamente. O que se justifica, ainda, nos erros de planejamento de prazo (aditivos de prazo de finalização da obra) e projeções orçamentárias incorretas (aditivo de supressão de valor orientado pela área técnica da AP local).

Por fim, a partir dessa análise, outros fatos visíveis poderiam servir como características preditoras para possível melhoria da modelagem de risco, como: o número e tipo de aditivos implementados no contrato e distância da empresa ao local de execução do contrato.

9.2 Contrato II - Confecção de materiais gráficos

O segundo contrato explorado obteve 97% de chances de rescisão, teve início em 03/02/2017 e vigência até 31/12/2017. Com valor de R\$ 187.562,50, contemplou o objeto da *"Confecção de materiais gráficos destinados à manutenção das atividades das secretarias diversas e fundo municipal de Saúde de (...)".* Anterior à rescisão do contrato, originado por licitação na modalidade pregão presencial (onde, neste caso, não houve presença de outros concorrentes), foram empenhados R\$ 81.500,00 para pagamentos até 13/06/2017 e não houve nenhum termo aditivo de prazo ou valor. A empresa responsável iniciou suas atividades em dezembro de 2013, à época constava no cadastro RF como microempresa (que pode possuir um faturamento anual de até R\$ 360 mil) e um registro de endereço localizado a cerca de 49 quilômetros de distância da localidade da contratante.

De acordo com a decisão administrativa, *"O presente DISTRATO, se dá por motivo de Interesse Público, motivada por conveniência e oportunidade da Administração (...)"* [sic], seguido pela abertura de um procedimento de auditoria para averiguação do contrato, em que

conclui que: "(...) a auditoria entende estarem presentes o *fumus boni juris* (faturamento incompatível com a condição de microempresa; endereço coincidente com outras 03 PJ ativas; percentual de 'carona' de 297%, em desacordo com o estabelecido nos itens 9.8.2 e 9.8.3 do edital, que estabelece 100%) e do *periculum in mora* (contrato em execução, com pagamentos que já totalizam R\$ 81.500,00), requisitos necessários para a providência cautelar por parte deste Tribunal de Contas, prevista no artigo 195 do Regimento Interno, no sentido de suspender todos os atos decorrentes do Pregão Presencial nº 006/2017 – SRP, até ulterior pronunciamento deste Tribunal de Contas".

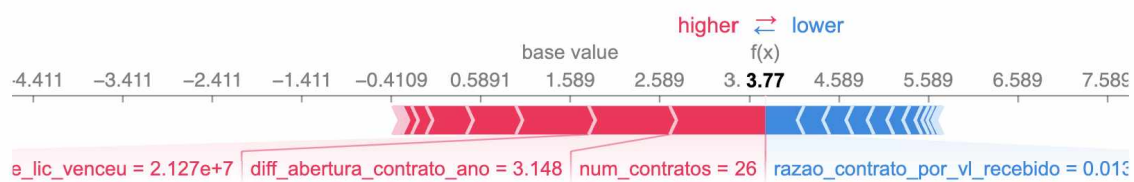


Figura 9.2: Composição do risco de rescisão – Contrato II

O contrato em questão representa 1% do valor total somado em contratos pela empresa com a AP, o que permite inferir, a partir do valor do contrato, que a empresa foi um fornecedor frequente e/ou firmou outros contratos de grande valor com a AP em seu tempo de atuação (três anos a contar desde a data de abertura da empresa e data de assinatura do presente contrato).

Utilizando o conjunto de dados de teste, ao considerar um subconjunto de empresas de até quatro anos e do mesmo porte, a média de participação em licitações foi de 22 (e mediana 10), a média de licitações vencidas foi 10 (e mediana 5) e a média de contratos firmados foi de 8 (e mediana 4). A empresa responsável pelo contrato, especificamente, participou de 74 licitações, venceu 28 delas e firmou 26 contratos (até a respectiva data de assinatura do documento). A partir dessa relação consecutiva e profícua, e dos valores destacados, é surpreendente que uma empresa relativamente jovem e de porte pequeno tenha tido um desempenho licitatório excepcional e uma soma de ganhos notável.

O grande acúmulo de contratos e do montante financeiro somado em licitações em relação à pouca idade da empresa, como detalhado anteriormente, foram comportamentos corretamente capturados pela modelagem através das características "número de contratos", "diferença entre a abertura da empresa e a data de assinatura do contrato" e "montante em

licitações vencidas pela empresa", como é possível observar na Figura 9.2. O "número de licitações nas quais a empresa participou" e o "total ganho em contratos" também foram características que contribuíram positivamente na composição do risco do contrato. Por outro lado, o atributo do porte da empresa, que também demonstra sua capacidade de força de trabalho, não foi uma característica relevante na composição do risco de rescisão.

Com base no detalhamento da auditoria, outras características possivelmente promissoras para experimentação em modelagens futuras são: o número de empresas concorrentes na licitação, caso faça sentido na modalidade, e endereços e/ou números de telefone coincidentes no registro de outras empresas. A primeira representa o possível comprometimento do caráter concorrencial do processo licitatório (resguardado por lei), enquanto a segunda denota um possível comportamento de empresas de fachada (que podem informar endereços físicos enganosos).

9.3 Contrato III - Fornecimento de alimentos

O terceiro contrato, com 98% de chances de rescisão, teve origem a partir de uma licitação na modalidade pregão presencial, tendo início no dia 16/03/2017, com vigência até 16/03/2018, e seu valor, de R\$ 198.600,00, contemplava a *"Aquisição de Gêneros Alimentícios (Farinha de Arroz e Farinha Láctea) para a atender ao Programa (...)"*. A empresa de micro porte, apesar de localizada na mesma cidade de fornecimento do objeto do contrato, foi aberta apenas nove meses antes da assinatura do contrato. Foi participante de uma única licitação, com uma proposta, e foi vencedora da mesma, um percentual de 100% de vitória.

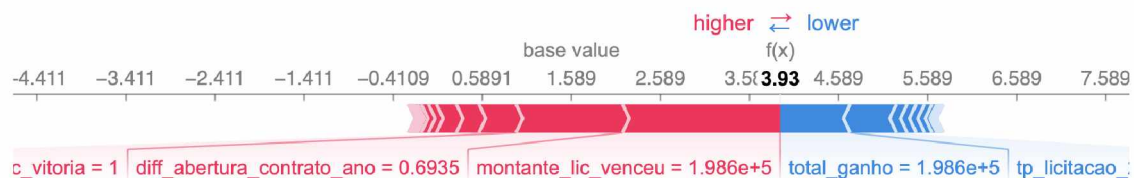


Figura 9.3: Composição do risco de rescisão – Contrato III

O modelo, por sua vez, foi capaz de capturar comportamentos frequentemente tidos como suspeitos, como um grande montante financeiro em licitações vencidas, associado a um porte micro (primeira e quinta características de maior contribuição na composição do risco), um

grande percentual de vitórias, além da pouca idade da empresa (terceira e segunda características de maior contribuição, respectivamente), como exibido na Figura 9.3.

A rescisão do contrato dá-se, entretanto, por solicitação da empresa contratada, que se justifica com *"Tal solicitação se dá em virtude do disposto no art. 78, X, da Lei 8.666/93: Art 78, tendo em vista que a empresa está em processo de dissolução de sociedade"*. Nesse sentido, apesar da identificação semanticamente coerente de características suspeitas de contrato e empresa contratada, a rescisão do contrato ocorreu por fatos cuja ocorrência não pode ser atribuída às características disponíveis no conjunto de dados utilizado para aprendizagem do modelo.

9.4 Contrato IV - Prestação de serviços terceirizados

Já em outro contrato, um comportamento semelhante ocorre em uma contratação de R\$ 2.237.233,00 (com 96% de chance de rescisão) cujo objetivo era *"Registro de preço para eventual contratação de empresa especializada na prestação de serviços terceirizados, de natureza contínua, para fornecimento de profissionais capacitados para atender as eventuais necessidades das Secretaria de Municipal de Educação, Secretaria Municipal de Desenvolvimento Social, Secretaria Municipal de Saúde, Secretaria Municipal de Infraestrutura e Secretaria Municipal de Serviços Públicos da Prefeitura Municipal (...)".*

Esse contrato, que teve início em 05/02/2018 e vigência até 04/02/2019, foi originado a partir de uma licitação na modalidade pregão presencial e era de responsabilidade de uma empresa sem enquadramento de porte (em que não há limite definido para seu faturamento, podendo ultrapassar R\$ 4,8 bilhões e ter mais de 100 funcionários). Ainda, localizada a uma distância de 303 quilômetros da unidade gestora contratante, a empresa foi participante em 70 licitações, entretanto o referido contrato representa 78% do total financeiro somado em contratos firmados – um contrato, consequentemente, significativo em sua arrecadação.

Na Figura 9.4 vê-se que o comportamento de participação em licitações frente ao número de vitórias são características de grande influência na composição do risco de rescisão desse contrato – que podem significar, por exemplo, empresas participantes laranja, utilizadas para simular competitividade na licitação. O que é contrastante, por exemplo, com o fato de o tipo de licitação (pregão presencial) ter uma contribuição negativa na composição do risco.



Figura 9.4: Composição do risco de rescisão – Contrato IV

As características comportamentais descritas acima não são, enfim, semanticamente coerentes com a justificativa de rescisão fornecida no processo, que ocorreu "*Considerando, as irregularidades no Procedimento Licitatório (...), constatadas através do processo TC no (...) e pelo parecer jurídico no (...) da Assessoria da Comissão de Licitação do Município (...)*" [sic]. No relatório técnico e nos documentos de defesa jurídica dos gestores, entretanto, o detalhamento da rescisão diz respeito à ausência de folhas de ponto devidamente assinadas pelos funcionários subcontratados e também de comprovantes referentes aos devidos pagamentos.

Assim como no contrato III, no contrato IV, os motivos diretos da rescisão não tiveram sua semântica concretamente ligada às características utilizadas para a modelagem do risco. Entretanto, outras características foram indicativas de risco relevantes e estão alinhados com os motivos de suspeição de especialistas. Como, por exemplo, um alto número de participações em licitações alinhado ao baixíssimo número de vitórias nessas licitações.

Essa empresa, em específico, fornece funcionários para serviços terceirizados e ao mesmo tempo é uma construtora. Essa atuação genérica, segundo especialistas, é um comportamento que pode levantar suspeitas de empresas de fachada que tentam fraudar o caráter de concorrência das licitações. Assim, o número de atividades econômicas de uma empresa pode ser aproveitado como característica preditora em modelagens futuras.

9.5 Contrato V - Construção de quadra escolar coberta

O objeto do quinto contrato analisado, vigente inicialmente por cerca de seis meses, fazia referência a "*Contratação de empresa do ramo de engenharia para construção de uma quadra coberta com vestiário, padrão FNDE anexa a escola (...)*" [sic], no valor de R\$ 507.450,00, e a tomada de preços foi a modalidade da licitação originária. No momento da firma do

contrato, a empresa responsável havia sido vencedora em quase 60% das licitações que participou, mesmo tendo sido aberta, no momento de início do contrato, há pouco mais de um ano.

Entretanto, mesmo com a pouca idade da empresa, grande frequência de vitórias e valor do alto do contrato, características que poderiam gerar alerta sobre a inexperiência ou capacidade técnica limitada da empresa – e que estão refletidas em preditores –, é possível verificar na Figura 9.5 que esses comportamentos não foram relevantes na composição do risco de rescisão desse contrato. Apesar da modalidade da licitação contribuir positivamente na composição do risco, as demais características que destacam-se positivamente em relação ao risco imprimem pouco impacto. A influência do PIB *per capita* do município onde o contrato foi firmado se sobressai em relação às demais características e é determinante para que o modelo classifique, erroneamente, o contrato como falso negativo.



Figura 9.5: Composição do risco de rescisão – Contrato V

A rescisão ocorreu por vontade de ambas as partes, devido ao descumprimento das obrigações contratuais pela contratada em relação à entrega do objeto do contrato no prazo acordado – apesar de não haver registro no Tramita acerca de aditivos de tempo solicitados pela empresa.

Ainda que o PIB *per capita* possa ocasionar melhora no desempenho do modelo, nesse caso, sua influência gerou viés na classificação, sobressaindo-se às características do próprio contrato e histórico da empresa. Nesse caso, é possível que a limitação da influência de determinadas características ou a definição de pesos pudesse mitigar a ocorrência dessa contradição.

9.6 Contrato VI - Instalação de itens de serralheria

Esse contrato teve início em 04/09/2017 e vigência até 31/12/2017, celebrando a "Contratação de empresa especializada para fornecimento e instalação de elementos de serralheria

para diversas secretarias deste município (...)" pelo valor de R\$ 163.400,00. O contrato foi classificado pelo modelo como de alto risco de rescisão, porém não possuía nenhuma mutação de interesse relacionada — um caso falso positivo.

No momento de sua assinatura esse contrato representava todo o faturamento da empresa com serviços à AP, sendo decorrente de uma licitação na modalidade pregão presencial sem nenhuma proposta concorrente. Foi o primeiro contrato da empresa ganhadora com a AP e teve duração de três meses. A empresa vencedora, por sua vez, possuía apenas seis meses desde sua abertura e porte de microempresa.

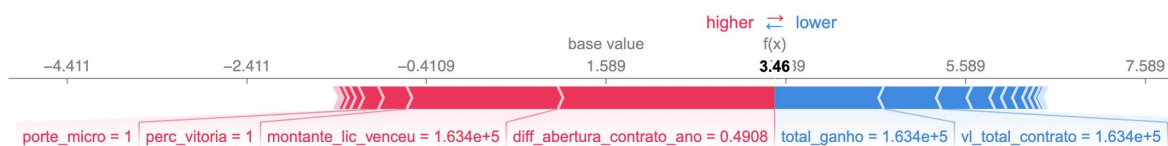


Figura 9.6: Composição do risco de rescisão – Contrato VI

Por ser classificado um contrato falso positivo, não há um parecer técnico da corte de contas com o qual seja possível comparar características relevantes para sua mutação. Entretanto, é possível analisar os fatos da própria contratação e compará-los com as características relevantes na composição do risco definidas pelo modelo e com características relevantes em outros contratos positivamente classificados.

A idade da empresa, seu porte e o montante financeiro vencido em licitações são características que contribuem positivamente na composição do risco de rescisão. Nos exemplos corretamente classificados analisado anteriormente, viu-se que a pouca idade e porte de microempresa são características que podem estar associadas à inexperiência com os processos públicos de prestação de serviço ou capacidades técnicas limitadas.

Como o conjunto de rótulos utilizado é provável de subnotificação, uma vez que o conjunto completo de contratos não foi inspecionado, esse tipo de caso pode ser, de fato, um falso positivo, ou um contrato de alto risco que não foi previamente analisado. Ainda, esse tipo de caso pode guiar esforços de análise em contratos que já foram encerrados e não receberam denúncias durante sua execução.

9.7 Discussão

Na literatura, o cálculo de métricas avaliativas (acurácia, precisão, revocação, por exemplo), utilizando a previsão do modelo no conjunto de dados de teste e a comparação do modelo com classificadores propostos anteriormente, são métodos comumente utilizados para realizar a avaliação do desempenho preditivo dos modelos de classificação de risco. A validação conceitual ou prática desses modelos, no entanto, nem sempre é realizada. Consultar especialistas para avaliar regras condicionais preditivas em classificadores bayesianos [14], ou comparar o desempenho do modelo com ferramentas mais simples [15] foram estratégias já aplicadas na prática para reforçar a validade do conhecimento extraído por meio de inferências e verificar a aplicabilidade dos classificadores propostos.

Na intenção de preencher a lacuna relativa à validação conceitual e prática das ferramentas propostas, neste capítulo buscou-se compreender se o comportamento de risco aprendido pelo modelo é correspondente ao constructo de risco de rescisão constituído pela variável resposta utilizada e se pode ser descrito através dos preditores. Utilizando relatórios técnicos relativos às próprias mutações contratuais, informação utilizada para determinar o rótulo de risco de rescisão, foi possível analisar a consonância entre as características preditivas relevantes na composição do risco para a classificação do contrato e seus motivos reais de rescisão.

Ainda, considerando que a contratação pública em cada um dos diferentes níveis de governo (federal, estadual e municipal) possui suas próprias características, é fundamental explorar como o risco de rescisão é materializado na prática da modelagem – nesta pesquisa, para contratos municipais brasileiros, nível ainda não abordado anteriormente. É importante ressaltar que a realização dessa análise exploratória de caráter inovador é possível, principalmente, pelo uso de uma variável dependente que denota um comportamento de risco específico e real do próprio contrato.

A partir da análise manual de uma amostra de quatro contratos verdadeiros positivos extraídos do conjunto de dados de teste, vê-se que os resultados das classificações possuem verossimilhança razoável com as características determinantes para a rescisão prática dos contratos.

Como observado, nos dois primeiros contratos analisados, o contexto semântico das ca-

racterísticas preditoras relevantes na composição do risco pode ser reconhecida nas justificativas de rescisão contratual publicada pela própria AP ou auditores da corte de contas. O que corrobora não apenas a relevância do conjunto de descritores de comportamentos de risco utilizados como características preditivas, mas também o aprendizado do risco específico de rescisão contratual. Por outro lado, nos dois contratos analisados em seguida, as características relevantes na composição do risco não possuíam semântica diretamente relacionada aos motivos das rescisões contratuais – ainda que fossem indicativos coerentes de suspeição de risco. Em qualquer modelo de AM, o aprendizado e as previsões são limitados pelas características e observações utilizadas como fonte de conhecimento. Assim, esta análise busca auxiliar na compreensão do comportamento do modelo, sem a pretensão de determinar as causas exatas do comportamento de alto risco.

Ainda, a análise de um contrato falso negativo e um falso positivo revelou a necessidade de melhorias específicas na modelagem. No primeiro caso, a limitação da influência de determinadas características pode ser necessária para mitigar vieses que possam ofuscar outros indicativos de risco relevantes. Já no caso falso positivo, a presença de características relevantes associadas ao alto risco de rescisão pode indicar contratos já encerrados mas com possíveis irregularidades.

Por fim, é importante destacar que as características utilizadas foram baseadas em indicadores de risco presentes na literatura ou já conhecidos pelos praticantes da área e, como visto, não representam a totalidade dos comportamentos suspeitos. Em vista disso, de modo a contemplar a dinamicidade na atuação das empresas contratadas quanto ao histórico característico de risco e manter o modelo atualizado, é necessário um esforço contínuo para refletir novos comportamentos em características preditoras.

Capítulo 10

Conclusões

A possibilidade de estimação do risco de rescisão de contratos públicos de maneira antecipada por meio do uso de AM, de características comportamentais das empresas e dos próprios contratos representa uma oportunidade singular para otimização do uso de recursos humanos e financeiros no Brasil.

A princípio, quantificou-se o desempenho de uma matriz de risco utilizada na prática para estimação do risco de empresas em contratação com a AP. A ampliação dessa avaliação para os contratos sob responsabilidade das empresas corretamente classificadas pela ferramenta fornece evidências para enquadrar o problema da estimação de risco em um nível mais acionável: o do contrato. Na prática, esse é o nível mínimo de vínculo entre a administração pública e as empresas contratadas. Nesse sentido, constatou-se que esforços anteriores para modelagem e estimação de risco de contratos públicos, até então, utilizaram rótulos em nível de empresa para a classificação de risco. Isto significa que, ainda que a previsão do modelo gere uma probabilidade de classificação relativa ao contrato, na prática, a variável resposta, que denota o risco, não representa um comportamento específico daquela observação, senão uma generalização do risco ali apontado.

A modelagem do risco de rescisão contratual foi, então, realizada a partir de um conjunto de características preditoras que abrangem o próprio contrato, a empresa contratada e seu histórico de relacionamento com a AP. Além disso, utilizou-se um conjunto inédito e representativo de contratos nos quais irregularidades ou falhas contratuais foram constatadas, servindo como variável resposta. Considerando o pior cenário de avaliação da performance do modelo proposto – limite inferior do IC com 95% de confiança –, treinado com o algo-

ritmo XGBoost, a performance do modelo proposto para a AUC-PR é sete vezes superior a um classificador aleatório para o conjunto dados utilizados. Esse resultado fornece evidência para abordar o problema da estimação de risco da contratação em nível da entidade contrato – e não de empresa. Na prática, o uso desse modelo (*SAGRES RF*) permitiria o acompanhamento adequado de cerca de 7% (R\$ 12 milhões) do montante financeiro total sabido como alto risco (R\$ 196 milhões).

Em seguida, considerando o cenário de severo desbalanceamento entre as classes de risco em que a classe de interesse representa 1,5% do conjunto de dados disponível, foram experimentadas estratégias de balanceamento artificial entre classes e comparadas com o aprendizado sensível ao custo. Para o contexto de contratos municipais, os modelos treinados com conjuntos de dados balanceados não demonstraram melhora estatisticamente significativa em sua performance preditiva em relação ao treinamento com pesos definidos para cada classe de risco. Esse resultado sugere que a mera replicação das observações não produz padrões relevantes ao aprendizado acerca da classe minoritária pelo modelo, enquanto a diminuição da classe majoritária resulta em perda de informações relevantes que auxiliam na diferenciação de ambas as classes.

Em complemento, a modelagem foi refinada por meio da experimentação de novos preditores de risco e do uso de pesos relacionados ao valor financeiro do contrato. No primeiro esforço, características latentes acerca do tema objeto do contrato (extraídas utilizando LDA) e a característica econômica do PIB *per capita* da cidade na qual o contrato foi firmado foram adicionadas individualmente ao conjunto de características inicial. Os resultados apontam que o uso de características acerca do tema do contrato impacta de maneira marginal a performance preditiva do modelo, apesar de tornar sua classificação mais conservadora quando comparada a do modelo *SAGRES RF*. Por outro lado, a adição do PIB *per capita* confere ao modelo uma maior capacidade de distinguir contratos da classe de alto risco e torna plausível sua superioridade de performance em relação ao modelo *SAGRES RF*. Contudo, não há evidência de diferença estatisticamente significativa ao comparar *SAGRES RF Tópicos* e *SAGRES RF PIB percapita* com o modelo *SAGRES RF*.

Financeiramente, na prática, o modelo *SAGRES RF Tópicos* representa maior desvantagem na identificação de contratos corretamente classificados como alto risco de rescisão, identificando apenas R\$ 6,2 milhões. Já o modelo *SAGRES RF PIB percapita* representa o

melhor custo-benefício até então, permitindo o acompanhamento apropriado de R\$ 74 milhões de reais, e reduzindo os montantes classificados em falsos positivos (R\$ 56 milhões) e falsos negativos (R\$ 122 milhões).

Ainda no sentido de assegurar o máximo retorno financeiro das investidas de acompanhamento da execução dos contratos, um conjunto de pesos definidos com base no valor de cada contrato foi adicionado à etapa de treinamento. Dessa maneira, contratos de maior valor teriam um custo de erro maior, resultando em um modelo que erra menos a classificação de contratos de grande porte financeiro. O IC da diferença entre esse e o modelo *SAGRES RF PIB per capita* não indica diferença estatisticamente significativa entre os modelos. Entretanto, a previsão de risco dos contratos indica um aumento no montante financeiro de contratos de alto risco corretamente classificados (R\$ 84 milhões), bem como diminuição do montante de falsos positivos (R\$ 85 milhões) e falsos negativos (R\$ 111 milhões).

Dentre os conjuntos de características construídas a partir de fontes de dados públicas (SAGRES, RF e Tópicos), é provável que a ausência do conjunto de tópicos não resulte em diminuição da capacidade preditiva do modelo. Por outro lado, é plausível que a ausência de características provenientes do SAGRES tenha impacto negativo na performance do classificador. Esse resultado pode auxiliar praticantes futuros na obtenção de conjuntos de dados prioritários para a modelagem de risco de contratos locais.

Ainda, o comportamento do risco de rescisão foi explicado a partir das previsões individuais do modelo *SAGRES RF PIB per capita* para o conjunto de dados de teste e a partir da distribuição dos valores características preditoras utilizadas. O panorama trazido nessa análise representa um avanço no entendimento do comportamento do risco de rescisão contratual e pode aumentar a confiança na aplicação do modelo em cenários práticos à medida que torna sua previsão mais palpável. Ao filtrar apenas os contratos corretamente classificados pelo modelo, uma segunda análise de explicabilidade é realizada com o objetivo de compreender, efetivamente, como o risco foi composto nos casos verdadeiramente positivos. Os resultados dessa análise evidenciam as características mais relevantes na composição do risco de rescisão e ainda sua ambiguidade de contribuições.

Por fim, considerando contratos classificados com maior chance de rescisão no conjunto de dados de teste, analisou-se o alinhamento do comportamento de aprendizado do modelo com pareceres técnicos de justificativas das rescisões publicados pelo TCE-PB. Com isso,

verificou-se um alinhamento entre a influência das características preditoras de risco propostas na modelagem e comportamentos reconhecidamente irregulares por especialistas na área. Com base nesses resultados, é fundamental pontuar que a relevância dessa análise está diretamente relacionada à validação do constructo de risco proposto nesse trabalho, à geração de credibilidade na previsão do modelo, ao aperfeiçoamento da governança da contratação pública, bem como à aplicação prática desse modelo.

Assim como em outros estudos, a proposição de modelos de estimação de risco de contratos nesse trabalho não almeja substituir as ferramentas de controle atuais, senão complementá-las, em um esforço proativo para maximizar a eficiência dos recursos disponíveis. Com isso, a partir dos resultados apresentados, essa tese contribui, principalmente, em três perspectivas para o avanço da estimação de risco de rescisão em contratos públicos brasileiros por meio da aplicação de AM.

Primeiramente, cria-se evidência para o enquadramento da estimação do risco em nível de contrato, o que oportuniza a evolução da estimação de risco contratual em municípios de outros estados e países. Em segundo lugar, agências governamentais ou da sociedade civil (que podem possuir distintas condições de controle) podem valer-se das indicações de risco para direcionar esforços de acompanhamento ou fiscalização de contratos específicos. Por último, valida a modelagem do constructo de risco a partir das características propostas e variável resposta utilizada. As análises podem, ainda, apontar diretrizes para reformulações em processos de governança nas contratações públicas e servir de inspiração para a proposição de novas características preditivas.

10.1 Limitações

As principais limitações relacionadas a essa pesquisa foram descritas ao longo do documento, em seções específicas. Contudo, é válido destacar que o uso de um conjunto de rótulos de rescisão contratual construído manualmente a partir de auditorias e sanções anteriores pode, porventura, conter o viés pessoal e inerente ao modo de trabalho de cada servidor responsável.

Ainda, a representação do contrato, nesse trabalho materializada através de características estáticas, não captura a dinamicidade do processo de contratação — ou mesmo da

execução do contrato, uma vez que não estão inclusos dados sobre termos aditivos (de tempo ou dinheiro) e tampouco dados acerca da execução orçamentária (empenhos e pagamentos).

Finalmente, diferentes esferas de governo terão características próprias em relação ao desembolso de recursos e características contratuais. A replicação desse trabalho para diferentes contextos está condicionada à análise e validação de seus resultados.

10.2 Direções futuras

Esta tese explorou o uso e impacto de novas ou já conhecidas características preditoras no comportamento de risco. Nesse sentido, novas características foram propostas no Capítulo 9, a partir das decisões técnicas do TCE-PB. Em adição, a partir da literatura em outras áreas, outras características podem ser utilizadas, como: indicadores sobre o nível de abertura de dados de cada município; características históricas sobre cada unidade gestora e sobre a experiência dos funcionários responsáveis pelo acompanhamento dos contratos; além da segregação entre os anos de experiência da empresa em contratação com o setor público e anos de experiência em atividade. Contudo, um esforço para geração de um conjunto de rótulos de maior tamanho e abrangência é também fundamental para contemplar a diversidade de comportamentos de risco presentes na prática da contratação.

A experimentação de modelos de linguagem para geração de *embeddings* a partir das descrições dos objetos contratados pode aprimorar a tarefa de identificação da área do contrato. Essa estratégia pode contribuir para mitigar ambiguidades semânticas nas descrições dos objetos contratuais, melhorando a precisão na identificação dos temas de cada contrato e tornando a análise de contexto mais rica e profunda. Essa abordagem é especialmente relevante devido ao bom desempenho desses modelos em tarefas semelhantes.

Por fim, a validação externa do modelo por especialistas, por meio da previsão de risco para contratos vigentes, pode ser um desdobramento futuro dos resultados aqui apresentados.

Bibliografia

- [1] Victor Aguiar Jardim de Amorim. Licitações e contratos administrativos: teoria e jurisprudência. *Coordenação de Edições Técnicas*, 2017.
- [2] Samira e Mouhoub Malek Anowar, Farzana e Sadaoui. Auction fraud classification based on clustering and sampling techniques. In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 366–371. IEEE, 2018.
- [3] Pierre e Mazer Emmanuel e Cobbe Paulo Balaniuk, Remis e Bessiere. Risk based government audit planning using naïve bayes classifiers. In *Advances in Knowledge-Based and Intelligent Information and Engineering Systems*, 2012.
- [4] David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022, 2003.
- [5] Giuseppe Bonaccorso. *Machine learning algorithms*. Packt Publishing Ltd, 2017.
- [6] Francesco Bonelli and Sandro Cabral. Efeitos das competências no desempenho de contratos de serviços no setor público. *Revista de Administração Contemporânea*, 22(4):487–509, 2018.
- [7] Kendrick Boyd, Kevin H Eng, and C David Page. Area under the precision-recall curve: point estimates and confidence intervals. In *Joint European conference on machine learning and knowledge discovery in databases*, pages 451–466. Springer, 2013.
- [8] Brasil. Lei nº8.666, de 21 de junho de 1993. regulamenta o art. 37, inciso XXI, da Constituição Federal, institui normas para licitada administração ccões e copública e dá outras providências. *Diário Oficial União*, 1993.

- [9] Brasil. Decreto nº9.203, de 22 de novembro de 2017. dispõe sobre a política de governança da administração pública federal direta, autárquica e fundacional. *Diário Oficial União*, 2017.
- [10] Brasil. Lei nº14.133, de 1 de abril de 2021. regulamenta o art. 37, inciso XXI, da Constituição Federal, institui normas para licitada administração ccões e copública e dá outras providências. *Diário Oficial União*, 2021.
- [11] Sandro Cabral. Reconciling conflicting policy objectives in public contracting: The enabling role of capabilities. *Journal of Management Studies*, 54(6):823–853, 2017.
- [12] Leonardo J e Da Rocha Henrique A e Mendes Gilson L Carvalho, Rommel N e Sales. Using bayesian networks to identify and prevent split purchases in brazil. In *Proceedings of the Eleventh UAI Conference on Bayesian Modeling Applications Workshop-Volume 1218*, pages 70–78. CEUR-WS. org, 2014.
- [13] Rommel Carvalho, Eduardo de Paiva, Henrique da Rocha, and Gilson Mendes. Using clustering and text mining to create a reference price database. *Learning and NonLinear Models*, 12(2014):38–52, 2014.
- [14] Rommel Novaes Carvalho, Ricardo Silva e Carvalho. Bayesian models to assess risk of corruption of federal management units. In *BMA@ UAI*, pages 28–35, 2016.
- [15] Rommel Novaes e Ladeira Marcelo e Monteiro Fernando Mendes e de Oliveira Mendes Gilson Libório Carvalho, Ricardo Silva e Carvalho. Using political party affiliation data to measure civil servants’ risk of corruption. In *2014 Brazilian Conference on Intelligent Systems*, pages 166–171. IEEE, 2014.
- [16] Kevin W e Hall Lawrence O e Kegelmeyer W Philip Chawla, Nitesh V e Bowyer. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16:321–357, 2002.
- [17] Olivier Compte, Ariane Lambert-Mogiliansky, and Thierry Verdier. Corruption and competition in procurement auctions. *Rand Journal of Economics*, pages 1–15, 2005.
- [18] Maurizio Conti, Pier Luigi e Naldi. Detection of anomalous bids in procurement auctions. *Decision Support Systems*, 46(1):420–428, 2008.

-
- [19] Geoff Cumming and Sue Finch. Inference by eye: confidence intervals and how to read pictures of data. *American psychologist*, 60(2):170, 2005.
- [20] Brasil Tribunal de Contas da União. Licitações e contratos: orientações e jurisprudência do tcu. *Secretaria Especial de Editoração e Publicações*, 2010.
- [21] Tribunal de Contas da União. Referencial de combate a fraude e corrupção: Aplicável a órgãos e entidades da administração pública. *Brasília: Secretaria de Métodos e Suporte ao Controle Externo (Semec)*, 2017.
- [22] Jesse Davis and Mark Goadrich. The relationship between precision-recall and roc curves. In *Proceedings of the 23rd international conference on Machine learning*, pages 233–240, 2006.
- [23] Mauricio Barros de Jesus, Gladston Luiz da Silva, Marcelo Ladeira, and Gustavo CG Van Erven. Using text mining to categorize the purpose of public spending for the benefit of transparency and accountability. In *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)*, pages 263–267. IEEE, 2019.
- [24] Órion Darshan Winter De Lima and Nazareno Andrade. Fairness in risk estimation of brazilian public contracts. pages 57–64, 2019.
- [25] Joselito Mendes de Sousa Júnior, Vinícius Ponte Machado, Rodrigo de Melo Souza Veras, Roney Lira de Sales Santos, Bruno Vicente Alves de Lima, Aline Montenegro Leal Silva, Francisco Alysson da Silva Sousa, and Francisco das Chagas Imperes Filho. Public administration suppliers classification model based on supervised machine learning. In *Proceedings of the XVIII Brazilian Symposium on Information Systems*, pages 1–8, 2022.
- [26] Rommel N e Carvalho Ricardo S e Ramos Guilherme N Domingos, Silvio L e Carvalho. Identifying it purchases anomalies in the brazilian government procurement system using deep learning. In *2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 722–727. IEEE, 2016.
- [27] Jorge Gallego, Gonzalo Rivero, and Juan Martínez. Preventing rather than punishing:

- An early warning model of malfeasance in public procurement. *International Journal of Forecasting*, 37(1):360–377, 2021.
- [28] Manuel J García Rodríguez, Vicente Rodríguez Montequín, Francisco Ortega Fernández, and Joaquín M Villanueva Balsera. Public procurement announcements in spain: regulations, data analysis, and award price estimator using machine learning. *Complexity*, 2019, 2019.
- [29] Thiago Alencar Gomes, Rommel Novaes Carvalho, and Ricardo Silva Carvalho. Identifying anomalies in parliamentary expenditures of brazilian chamber of deputies with deep autoencoders. In *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 940–943. IEEE, 2017.
- [30] Ankit e Redmiles Elissa e Ghani Rayid Grace, Emily e Rai. Detecting fraud, corruption, and collusion in international development contracts: The design of a proof-of-concept automated system. In *2016 IEEE International Conference on Big Data (Big Data)*, pages 1444–1453. IEEE, 2016.
- [31] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*, volume II. Springer Science+Business Media, New York, USA, 2009.
- [32] Bruno M Henrique, Vinicius A Sobreiro, and Herbert Kimura. Contracting in brazilian public administration: A machine learning approach. *Expert Systems*, 37(5):e12550, 2020.
- [33] Matthew Hoffman, Francis Bach, and David Blei. Online learning for latent dirichlet allocation. *advances in neural information processing systems*, 23, 2010.
- [34] Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, et al. *An introduction to statistical learning*, volume 112. Springer, 2013.
- [35] Ó LIMA, Órion Darshan Winter de LIMA, et al. Justiça em aprendizagem de máquina na estimativa de risco de contratos públicos. 2020.
- [36] Scott Lundberg. A unified approach to interpreting model predictions. *arXiv preprint arXiv:1705.07874*, 2017.

- [37] Talita Lôbo de Menezes et al. Eficácia no uso de aprendizagem de máquina para estimação de risco em contratos públicos e empresas. 2019.
- [38] Tim Miller. Explanation in artificial intelligence: Insights from the social sciences. *Artificial intelligence*, 267:1–38, 2019.
- [39] Christoph Molnar, Giuseppe Casalicchio, and Bernd Bischl. Interpretable machine learning—a brief history, state-of-the-art and challenges. In *Joint European conference on machine learning and knowledge discovery in databases*, pages 417–431. Springer, 2020.
- [40] Toshio Mukai. *O Novo Estatuto Jurídico Das Licitações E Contratos Públicos: Comentários À Lei 8,666/93: com as Alterações Promovidas Pela Lei 8.883/94*. Editora Revista dos Tribunais, 1994.
- [41] Krystyna Napierala and Jerzy Stefanowski. Types of minority class examples and their influence on learning classifiers from imbalanced data. *Journal of Intelligent Information Systems*, 46:563–597, 2016.
- [42] Milena I Neshkova and Vaiva Kalesnikaite. Corruption and citizen participation in local government: Evidence from latin america. *Governance*, 32(4):677–693, 2019.
- [43] Mariana Peneda, José Cruz, and Pedro Sousa. Risk of corruption in public procurement in portuguese municipalities. how is it related to politics, gender of the mayor, and socioeconomic municipalities characteristics? *Public Integrity*, pages 1–13, 2024.
- [44] Maria Sylvia Zanella Di Pietro. Direito administrativo. In *Direito administrativo*, volume 27. Editora Atlas, 2014.
- [45] Hércules Antonio e Fernalda Edilson e Balaniuk Remis Rebouças, Rômulo Roseo e do Prado. Detecção de figurantes em pregões eletrônicos do governo federal brasileiro. *Informação & Tecnologia*, 2(2):5–21, 2017.
- [46] PR Reis and S Cabral. Para além dos preços contratados: fatores determinantes da celeridade nas entregas de compras públicas eletrônicas. *Revista de Administração Pública* 52, 107-125, 2018.

- [47] Leonardo J Sales. Risk prevention of public procurement in the brazilian government using credit scoring. Technical report, Observatório de Economia e Gestão de Fraude, 2013.
- [48] Rommel Novaes Sales, Leonardo Jorge e Carvalho. Measuring the risk of public contracts using bayesian classifiers. In *BMA@ UAI*, pages 7–13, 2016.
- [49] Ed Wilson Fernandes de Santana. *A utilização da informação contábil como um instrumento de controle social dos municípios paraibanos com mais de 50.000 habitantes*. PhD thesis, Dissertação (Mestrado em Contabilidade)-UnB, UFPB, UFRN e UFPE), Brasília, 2008.
- [50] Anwar Shah. Measuring government performance in the delivery of public services. *Handbook on Public Sector Performance Reviews*, 2, 2003.
- [51] Ravid Shwartz-Ziv and Amitai Armon. Tabular data: Deep learning is not all you need. *Information Fusion*, 81:84–90, 2022.
- [52] Leonardo J Sun, Ting e Sales. Predicting public procurement irregularity: An application of neural networks. *Journal of Emerging Technologies in Accounting*, 15(1):141–154, 2018.
- [53] István János Tóth and Miklós Hajdu. Factors affecting the corruption risk and intensity of competition in public procurement at the level of local government, 2020.
- [54] Ruben van den Goorbergh, Maarten van Smeden, Dirk Timmerman, and Ben Van Calster. The harm of class imbalance corrections for risk prediction models: illustration and simulation using logistic regression. *Journal of the American Medical Informatics Association*, 2022.
- [55] Rafael B Velasco, Igor Carpanese, Ruben Interian, Octavio CG Paulo Neto, and Celso C Ribeiro. A decision support system for fraud detection in public procurement. *International Transactions in Operational Research*, 28(1):27–47, 2021.
- [56] Alfredo Vellido, José David Martín-Guerrero, and Paulo JG Lisboa. Making machine learning models interpretable. In *ESANN*, volume 12, pages 163–172. Citeseer, 2012.

-
- [57] Maumita West, Jarrod e Bhattacharya. Intelligent financial fraud detection: a comprehensive review. *Computers & security*, 57:47–66, 2016.

Resumo - Trabalhos Anteriores

123

Ref.	Entidade	Dados	Contexto	Modelagem	Balanceamento	Métrica
2021 [27]	Contratos	SECOP, empresas investigadas ou multadas	Colômbia (Nível Nacional e Municipal)	Lasso, Gradient Boosting	Oversampling e Pesos	AUC-ROC
Continua na próxima página.						

2020 [32]	Contratos	ComprasNet, SI-CAF, RAIS, RF, TSE, CEIS	Brasil (Nível Federal)	Regressão Logística, KNN, Análise Discriminante, SVM, Floresta Aleatória	Undersampling	Erro tipo II (FN), AUC-ROC, Acurácia média
2016 [30]	Contratos	Banco Mundial	Contratos de desenvolvimento internacional	Floresta Aleatória, Regressão Logística, AdaBoost, SVM, GBC e KNN	-	Precisão
2018 [52]	Empresas	SIAFI, SIASG, RF, RAIS, SIAPE, TSE, CEIS	Brasil (Nível Federal)	Redes Neurais Profundas, Regressão Logística, Análise Discriminante	Oversampling aleatório	Acurácia, Precisão, AUC-ROC, F-scores
2016 [48]	Empresas	SIASG, CEIS	Brasil (Nível Federal)	Naive Bayes, Tree-Augmented Naive Bayes	Undersampling aleatório	Acurácia
2013 [47]	Empresas	SIASG, CEIS	Brasil (Federal)	Regressão Logística, Árvore de Decisão, Análise Discriminante	Undersampling aleatório	Acurácia
Continua na próxima página.						

2020 [24]	Empresas e Contratos	CGU, CEIS, SA-GRES	Brasil (Nível Federal e Municipal)	Regressão Logística, Floresta Aleatória, Rede Neural, SVM, KNN	Under e Over-sampling aleatórios, SMOTE	F1 score
-----------	----------------------	--------------------	------------------------------------	--	---	----------

Tabela A.2: Visão Geral da Literatura sobre Empresas e Contratos: Resultados.

Ref.	Baseline	Conclusão
2021 [27]	Classificador aleatório (AUC-ROC 0,5)	Os modelos produzidos são consideravelmente melhores do que o baseline definido pela métrica. AUC-ROC = 0,912.
2020 [32]	Regressão logística, classificador aleatório (AUC-ROC 0,5)	O modelo utilizando KNN foi superior aos demais. Taxa Erro Tipo II = 4,25% (contra 30,99% da Regressão Logística).
2016 [30]	Comparação entre modelos propostos	A precisão do GBC na configuração considerada pelo modelo é de 70%.
2018 [52]	Classificador aleatório (AUC-ROC 0,5) e comparação entre modelos propostos	O modelo utilizando Redes Neurais Profundas obteve o melhor desempenho. AUC-ROC = 0,878.
2016 [48]	Comparação entre algoritmos e trabalhos anteriores do autor	Não houve diferença estatística significativa entre ambos os modelos considerados.
2013 [47]	Press' Q Test para mensurar a significância do modelo em relação a um palpite aleatório, comparação entre modelos	A principal contribuição, citada pelos autores, foi a proposição e entendimento de <i>features</i> preditoras do risco das empresas. Os modelos gerados foram ligeiramente melhores que o baseline.
2020 [24]	Comparação entre modelos	Apesar do foco ser a análise de justiça, os modelos de risco mais promissores utilizaram Floresta Aleatória com balanceamento via SMOTE.

Apêndice B

Contratos por Tópico

Tabela B.1: Termos frequentemente associadas a cada tópico.

Tópico	Termos associados
Tópico 1	saude, medico_hospitalar, contratacao, veiculo, exames, infraestrutura, ministério, transporte_escolar, pereciveis, familia, transporte, ambulancia, casa, especificacoes_tecnicas, transito
Tópico 2	saude, hospital, odontologico, policlinica, area, postos, samu, caps, implantacao, odontologicos, frota_veicular, tecnicos, contratacao, oleo_diesel, gestao
Tópico 3	veiculos, frota, combustiveis, contratacao, medicamentos_farmacia, pneus, pneus_camaras, locados, combustiveis_lubrificantes, protetores, maquinas, maquinas_pesadas, filtros, vinculados, farmaceutica
Tópico 4	veiculos, locacao, servicos, pecas, contratacao, manutencao, frota, maquinas, veiculos_pesados, motorista, unico, pintura, pac, acessorios, escolar_ore
Tópico 5	exercício, aquisicoes, pecas_acessorios, ate, pacientes, judiciais, formacao, contratacao, caminhos, didatico, distribuicao, administracao_ate, tecidos, comunidades_rurais, passagens
Tópico 6	precos, atencao, medico, medicos, orgaos, pecas_automotivas, eletrico_hidraulico, psf, acoes, clinica, ofertado, continuados, governo, medicamentos_eticos, genericos_similares
Continua na próxima página.	

Tópico 7	contratacao, servicos, execucao, ruas, pregao, pavimentacao_paralelepipedos, pavimentacao_paralelepipedo, caminhao, distrito, domiciliares, projetada, licitacao_modalidade, aula, limpeza, residuos_solidos
Tópico 8	contratacao, servicos, engenharia, servico, recuperacao, informatica, execucao_obra, praca, execucao, pavimentacao_ruas, prestacao, upa, movel, vinte, festividades
Tópico 9	joao, sao, lote, contratacao, diesel, reposicao_pecas, basico, combustivel, lubrificantes, petroleo, compromisso, licitacao, camara, poliesportiva, hidraulico
Tópico 10	saude, medicamentos, contratacao, populacao, pavimentacao_drenagem, operacionais, judicial, base, mobiliario, urgencia, fms, tudo, predios, participacao, genero_alimenticio
Tópico 11	juridica, contratacao, equipamentos_informatica, combustiveis_petroleo, locacao, patrulha_mecanizada, propriedade, eventos, usuarios, esperanca, trecho, posto, som, instrumento, ventura_especificacoes
Tópico 12	especificacoes, instalacao, execucao_obras, comunidade, cras, onibus, creche, administrativas, escolares, sao_partes, peti, baraunas_exercicio, portas, utensilios, condicoes
Tópico 13	contratacao, saude, obra, solicitacao, confeccao, exames_laboratoriais, oficial, medicamentos_psicotropicos, pedra, grafico, alunos_ensino, comodato, urbano, fornecidos, fardamento
Tópico 14	generos_alimenticios, programas, limpeza, merenda_escolar, contratacao, construcao_quadra, fundamental, mobiliarios, creche, alimentacao_escolar, vestuario, escolar_coberta, escola_ensino, pavimentacao, refeicoes
Tópico 15	prestacao_servicos, contratacao, reforma, servicos, escola, execucao, insumos, sao_joao, ampliacao, peixe, mecanicos, esgotamento_sanitario, global, selecao_vantajosa, cujo_criterio
Continua na próxima página.	

Tópico 16	construcao, contratacao, agua, conclusao, convenio, ubs, limpeza_higiene, descartaveis, pessoal, nova, disposicoes, implantacao, trabalho, hidraulicos, predios_publicos
Tópico 17	equipamentos, saude, hospitalar, reforma_ampliacao, periodo, medicos_hospitales, contratacao, fundos, laboratorio, moveis_equipamentos, tratamento, entidades, materias, suporte, utilizacao
Tópico 18	farmacia, escolar, carnes, iluminacao, agricultura_familiar, educacao, injetaveis, fnde, medicamentos_basicos, creches, contratacao, pereci-veis_pereciveis, ensino_fundamental, veiculos_automotores, oficiais
Tópico 19	manutencao, preco, escolas, eletricos, contratacao, mao_obra, moveis, alunos, reposicao, farma, utensilios_domesticos, diaria, fardamentos, publicas, doacao
Tópico 20	educacao, expediente, administracao, ensino, eletrico, condicionado, hospita-les, graficos, acao, tratores, carro, pipa, laboratorial, quadra, caicara

Apêndice C

CNPJs Removidos

Tabela C.1: Lista de CNPJs de serviços essenciais.

CNPJ	Razão Social
09095183000140	Energisa Paraíba - Distribuidora de Energia S.A.
33000118001221	Telemar Norte Leste S/A.
33000118000179	Telemar Norte Leste S/A.
09123654000187	Companhia de Água e Esgotos da Paraíba Cagepa
01518579000141	A Uniao Superintendencia de Imprensa e Editora
09111832000150	Jornal Correio da Paraíba LTDA
04196645000100	Imprensa Nacional
00360305000104	Caixa Economica Federal
33530486000129	Empresa Brasileira de Telecomunicacoes S.A. Embratel
000000000000191	Banco do Brasil SA
60701190000104	Itau Unibanco S.A.
60746948184680	Banco Bradesco S.A.
28196889000143	Brasilseg Companhia de Seguros
90400888000142	Banco Santander (Brasil) S.A.
00058870164420	Folha de Pagamentos - Secretaria de Saúde
Continua na próxima página.	

33657248000189	Banco Nacional de Desenvolvimento Economico e Social / Sec de Infraestrutura
13937073000156	Paraíba - Secretaria da Fazenda
08907776000363	Polícia Militar do Estado - Pm/Pb

Apêndice D

Stopwords

'originais', 'atraves', 'adesao', 'aplicacao', 'nao', 'PB', 'Pb', 'pb', '-', '.', 'municipal', 'municipípio', 'empresa', 'atender', 'necessidade', 'necessidades', 'preço', 'parcelada', 'parceladas', 'parcelados', 'parcelado', 'parcela', 'prefeitura', 'deste', 'preços', 'prestação', 'aquisição', 'aquisições', 'registro', 'ontrtação', 'contratação', 'neste', 'nesta', 'deste', 'ad', 'ab', 'abc', 'abnt', 'abs', 'acc', 'supri', 'suprir', 'destinado', 'destinação', 'material', 'materiais', 'secretaria', 'secretarias', 'abaixo', 'turno', 'abastecida', 'básica', 'basica', 'pb', 'paraíba', 'paraiba', 'usuários', 'si', 'belém', 'conjunto', 'celebram', 'celebra', 'ii', 'iii', 'iv', 'uso', 'contrato', 'comum', 'cuitegi', 'tacima', 'queimadas', 'conforme', 'termo', 'termos', 'edital', 'referência', 'referencia', 'dia', 'especificação', 'especificações', 'anexo', 'anexos', 'constante', 'constantes', 'cajazeiras', 'campina', 'grande', 'santa', 'luzia', 'forma', 'localidades', 'localidade', 'local', 'locais', 'pública', 'públicas', 'público', 'públicos', 'publica', 'publico', 'lagoa', 'roça', 'bom', 'durante', 'diversas', 'diversos', 'destinados', 'destinadas', 'destinada', 'destinado', 'desta', 'outros', 'outras', 'eventual', 'eventuais', 'piancó', 'pertencentes', 'pertencente', 'janeiro', 'fevereiro', 'março', 'abril', 'maio', 'junho', 'julho', 'agosto', 'setembro', 'outubro', 'novembro', 'dezembro', 'cd', 'sendo', 'seja', 'disposição', 'municipalidade', 'josé', 'jose', 'unidade', 'bonfim', 'francisco', 'cidade', 'cidades', 'ser', 'serem', 'senhora', 'senhor', 'manaíra', 'maraira', 'realização', 'realização', 'realizacao', 'realizado', 'realizados', 'visando', 'teixeira', 'cabedelo', 'itaporanga', 'dias', 'forma', 'mamanguape', 'batista', 'sob', 'boa', 'empresas', 'maior', 'decorrente', 'decorrentes', 'visitas', 'visita', 'órgãos', 'executar', 'fazer', 'solicitação', 'solicitações', 'serviços', 'serviço', 'permanente', 'todas', 'toda', 'cacimbas', 'cavalos', 'gradual', 'pombal',

'igaracy', 'tipo', 'km', 'tipo', 'item', 'doze', 'dois', 'especializada', 'especializadas', 'especializado', 'especializados', 'geral', 'tribunal', 'atuando', 'jurupiranga', 'solânea', 'araruna', 'riachão', 'cacimba', 'número', 'nº', 'fará', 'cnpj', 'maria', 'vista', 'cariri', 'chamada', 'joão', 'pessoa', 'areia', 'pertinente', 'ano', 'ramo', 'administração', 'administrativa', 'sede', 'melhor', 'fim', 'demanda', 'itapororoca', 'exercício', 'exercícios', 'junto', 'atendimento', 'atendimentos', 'setor', 'setores', 'fundo', 'fornecimento', 'único', 'única', 'porte', 'um', 'dois', 'três', 'quatro', 'cinco', 'seis', 'sete', 'oito', 'nove', 'lucena', 'internacional', 'caçara', 'moura', 'bento', 'abrilhantar', 'seguinte', 'seguintes', 'guarabira', 'juru', 'alagoa', 'patos', 'art', 'valor', 'quixaba', 'seridó', 'centro', 'posterior', 'posteriores', 'dr', 'pedro', 'almeida', 'isabel', 'princesa', 'rita', 'rede', 'técnica', 'técnico', 'gestão', 'contas', 'apoio', 'resolução', 'alagoinha', 'prefeito', 'referente', 'tavares', 'piranhas', 'conceição', 'cajazeirinhas', 'pereira', 'exclusividade', 'exclusivo', 'fins', 'bernardino', 'através', 'costa', 'cpf', 'locação', 'fornecido', 'fornecer', 'aluguel', 'composição', '(s)', 'parte', 'integrante', 'integrantes', 'mesmo', 'mesmos', 'física', 'física', 'empresa(s)', 'empresa (s)', 'selecionar', 'edilidade', 'específico', 'específicos', 'núcleo', 'núcleos', 'rural', 'zona', 'credenciamento', 'catolé', 'gonçalves', 'sítio', 'adjacências', 'adjacencias', 'adjacência', 'adjacencia', 'autônomos', 'autônomas', 'autônomo', 'autônoma', 'autonomo', 'autonoma', 'campus', 'universidade', 'estadual', 'uepb', 'acesso', 'trajeto', 'trajetos', 'cachoeira', 'comercial', 'comercil', 'ltda', 'fogo', 'pedras', 'índios', 'indios', 'carnaúba', 'remoção', 'período', 'aberta', 'convocatório', 'controlados', 'controlado', 'controladas', 'controlada', 'situado', 'manhã', 'noite', 'demais', 'alto', 'todo', 'toda', 'aparecida', 'condições', 'condição', 'jerico', 'perante', 'organizações', 'social', 'sociais', 'organização', 'cabaceiras', 'terezinha', 'jurupiranga', 'picuí', 'possível', 'possivel', 'gerenciamento', 'interno', 'federal', 'federais', 'empreendedores', 'empreendedor', 'taperoá', 'compõem', 'conselho', 'capacidade', 'devendo', 'dever', 'entrega', 'entregar', 'pronto', 'sabadinho', 'horas', 'ações', 'seleção', 'esperança', 'prorrogar', 'horebe', 'onde', 'localizadas', 'localizada', 'localizado', 'localizados', 'elpídio', 'exigências', 'estabelecidas', 'estabelecida', 'brejo', 'santos', 'comprovação', 'superintendência', 'atividade', 'atividades', 'félix', 'sousa', 'rio', 'ceo', 'sr', 'demandas', 'demanda', 'zero', 'cruz', 'edgley', 'prestar', 'via', 'quantidade', 'quantidades', 'partir', 'anual', 'alterações', 'utilizados', 'utilizado', 'utilizadas', 'utilizada', 'olinda', 'exigidos', 'exigido', 'exigidas', 'exigida', 'nacional', 'rais', 'convênios', 'convênio', 'semana',

'semanas', 'fundação', 'qualidade', 'oliveira', 'mediante', 'kg', 'versa', 'decreto', 'tarde', 'execução', 'desenvolvimento', 'realizar', 'assistência', 'assistencia', 'manutenção', 'recuperação', 'estado', 'parecer', 'pareceres', 'bem', 'pessoas', 'tradicional', 'tradicionalis', 'bezerra', 'carvalho', 'jesus', 'brasil', 'inês', 'remigio', 'ocasião', 'instituto', 'poder', 'laudo', 'laudos', 'própria', 'próprias', 'próprio', 'próprios', 'preparo', 'mensal', 'cargo', 'ministério', 'condado', 'diretamente', 'dentro', 'meio', 'conclusão', 'assessoramento', 'duração', 'complementar', 'pré', 'pre', 'conde', 'profissional', 'profissionais', 'área', 'rua', 'avenida', 'documentos', 'documento', 'defesa', 'gabinete', 'coremas', 'atingida', 'atingido', 'atingidas', 'atingidos', 'paulista', 'tapada', 'lima', 'conformidade', 'pequenos', 'pequeno', 'notória', 'notórias', 'notório', 'notórios', 'abastecimento', 'acompanhamento', 'acordo', 'ar', 'civil', 'consumo', 'elaboração', 'câmara', 'lei', 'objeto', 'proposta', 'unidades', 'sistema', 'licitação', 'licitações', 'diário', 'diario', 'programa', 'segunda', 'atendendo', 'total', 'contratante', 'grupo', 'grupos', 'm³', 'mensais', 'bentinho', 'experiencia', 'experiência', 'âmbito', 'ambito', 'branco', 'regime', 'sitio', 'emancipação', 'política', 'politica', 'juninas', 'junina', 'combate', 'santana', 'meses', 'modelo', 'vargas', 'cândida', 'caráter', 'carentes', 'carente', 'processos', 'processo', 'enfrentamento', 'técnicos', 'consultoria', 'assessoria', 'plano', 'integrada', 'projeto', 'arquitetônico', 'militar', 'edson', 'ramalho', 'fazenda', 'marizópolis', 'humano', 'quais', 'qual', 'salgadinho', 'cuité', 'leilão', 'gênero', 'bonito', 'compra', 'alienação', 'regulamenta', 'regulamento', 'regulamentos', 'pregão', 'quantitativo', 'estabelecido', 'estabelecidos', 'planilha', 'orientação', 'chevrolet', 'direção', 'ingá', 'permanentes', 'passageiros', 'mg', 'objeto', 'representado', 'representada', 'representados', 'representadas', 'adesão', 'ata', 'final', 'monteiro', 'produzidos', 'produzido', 'produzida', 'produzidas', 'participação', 'funcionamento', 'volta', 'compreendendo', 'edição', 'pró', 'pro', 'imagem', 'ponto', 'reis', 'alves', 'periódica', 'departamento', 'legislativa', 'praça', 'apresentar', 'anos', 'santo', 'antonio', 'antônio', 'sebastião', 'barra', 'tabela', 'duas', 'fortalecimento', 'objetivando', 'servidores', 'atuação', 'urbana', 'juripiranga', 'presencial', 'incluindo', 'convivência', 'memorando', 'jurídica', 'conta', 'silva', 'soledade', 'chamamento', 'novo', 'venda', 'itens', 'básicas', 'basicas', 'população', 'emissão', 'emissões', 'assim', 'compor', 'corrente', 'objetivo', 'direta', 'litros', 'litro', 'leve', 'preventiva', 'corretiva', 'aplicação', 'leves', 'importância', 'passeio', 'reposição', 'domingos', 'todo', 'todos', 'velho', 'serra', 'letivo', 'necessários', 'necessário', 'corte', 'prestado', 'prestados', 'adminis-

trativo', 'vice', 'inclusive', 'menor', 'discriminado', 'sobre', 'uiraúna', 'aludido', 'remanescentes', 'remanescentes', 'utilitário', 'utilitários', 'produtos', 'produto', 'tendo', 'procedimento', 'procedimentos', 'mogeiro', 'derivados', 'derivado', 'recarga', 'novo', 'novos', 'futuro', 'futuros', 'futura', 'futuras', 'horário', 'horários', 'bananeiras', 'vicente', 'detentora', 'detentor', 'mil', 'repasse', 'cada', 'consta', 'mínima', 'ficar', 'desconto', 'básico', 'rodri-gues', 'atos', 'alusiva', 'alusivas', 'alusivo', 'alusivos', 'demandantes', 'demandante', 'remí-gio', 'riacho', 'garrotes', 'conservação', 'olho', 'desempenharem', 'sobreaviso', 'vencedora', 'vencedor', 'bayeux', 'especiais', 'bairro', 'municipais', 'nazarezinho', 'firmado', 'ma-mede', 'denominado', 'monte', 'branca', 'mínimo', 'minimo', 'equipar', 'ocorrer', 'ocorre', 'sexta', 'quinta', 'barbosa', 'gomes', 'contratada', 'interessados', 'interessadas', 'várzea', 'além', 'ida', 'preâmbulo', 'identificado', 'identificada', 'ocorrer', 'presente', 'gratuita', 'gratuito', 'percentual', 'loc'.