

## Modelo de classificação de licitações guiado por probabilidades para detecção de conluio

**David P. Galvão Júnior<sup>1\*</sup>, Gilberto F. de Sousa Filho<sup>1</sup>, Lucídio dos Anjos F. Cabral<sup>1</sup>, José Alysso Medeiros<sup>2</sup>, Thiago M. Albuquerque<sup>2</sup>, Marçal R. F. Lima Filho<sup>1</sup>, Sandro Marden Torres<sup>1</sup>**

<sup>1</sup> Universidade Federal da Paraíba, João Pessoa, PB

<sup>2</sup> Polícia Federal, João Pessoa, PB

\*Autor; e-mail: galvao2010@icloud.com

### RESUMO

Muitos trabalhos procuram analisar os valores financeiros das propostas de uma licitação pública, procurando classificá-la como fraude. Neste trabalho, propomos agrupar as empresas que participam das mesmas licitações fraudulentas, quantificando a probabilidade de cada grupo formar um conluio, e aplicar esta métrica na aprendizagem de máquina. Resultados demonstram uma melhora de até 9% às classificações da literatura.

**Palavras-chave:** agrupamento, licitação pública, aprendizado de máquina.

### Introdução

O Banco Mundial estima que no Brasil as licitações públicas representaram 20% do PIB de 2018 [Bosio 2020]. Por sua vez, fraudes em processos licitatórios adulteram a competição do processo, com fins ilícitos. A literatura busca detectar fraudes através de fatores como condições de mercado, preços das propostas e sua distribuição [Signor 2020]. Estes fatores foram utilizados para a classificação de licitações através do uso de aprendizagem de máquina [Rodríguez 2022].

### Objetivos

Neste trabalho, propomos agrupar as empresas por sua participação em licitações fraudulentas anteriores, e quantificar a probabilidade desta participação conjunta ser superior ao acaso.

### Métodos

Propomos o processo ilustrado na Figura 1 para classificação das licitações. Primeiro, agrupar as empresas pelo histórico de licitações (*dataset*). Depois, computar a probabilidade condicionada de uma licitação ser fraude usando o Teorema de Bayes. Por fim, construir um modelo de classificação das licitações guiado pelas probabilidades.



Figura 1. Modelo de detecção de fraudes.

### Resultados e Discussão

Para a comparação da métrica proposta neste trabalho e as métricas da literatura, criamos 4 configurações: (1) métricas de [Rodríguez 2022]; (2) métricas de [Signor 2020]; (3) probabilidade condicionada proposta e (4) originais das licitações.

Tabela 1. Resultados dos modelos de classificação de fraude.

|   |                  | Conjunto de Dados |         |        |         |
|---|------------------|-------------------|---------|--------|---------|
|   |                  | Brasil            | EUA     | Itália | Japão   |
| 1 | Melhor Algoritmo | SVM, K-NN         | SVM, NN | SVM    | K-NN    |
|   | Melhor Acurácia  | 95,00%            | 88,55%  | 82,14% | 94,44%  |
| 2 | Melhor Algoritmo | SVM, K-NN, NN     | -       | RF     | SVM, RF |
|   | Melhor Acurácia  | 85,00%            | -       | 83,93% | 94,44%  |
| 3 | Melhor Algoritmo | SVM, K-NN, RF, DT | NN      | RF     | SVM     |
|   | Melhor Acurácia  | 85,00%            | 97,43%  | 85,71% | 96,76%  |
| 4 | Melhor Algoritmo | SVM               | NN, RF  | RF     | SVM     |
|   | Melhor Acurácia  | 95,00%            | 88,79%  | 83,93% | 94,91%  |

Na Tabela 1 é possível observar que a Configuração 3 obtém a melhor acurácia de classificação em 3 das 4 bases de dados testadas.

### Conclusão

Agrupar empresas pelo seu histórico de licitações conjuntas apresentou robustez na classificação de fraude, comparada às métricas da literatura.

### Referências bibliográficas

- Bosio, E. and Djankov, S. (2020). How large is public procurement?
- Rodríguez, M. J. G., Rodríguez-Montequín, V., Ballesteros-Pérez, P., Love, P. E., and Signor, R. (2022). Collusion detection in public procurement auctions with machine learning algorithms. *Automation in Construction*.
- Signor, R., Love, P. E. D., Belarmino, A. T. N., and Olatunji, O. A. (2020). Detection of collusive tenders in infrastructure projects: Learning from operation car wash. *Journal of Construction Engineering and Management*.