

Otimização Inteligente de Contêineres no Porto de Santos: Integrando IoT e Machine Learning para Eficiência Operacional

Robert Richard das Neves Correia dos Santos

Fatec Baixada Santista - Rubens Lara

robert.santos@aluno.cps.sp.gov.br

ORCID: 0009-0001-8895-9712

José Augusto Theodosio Pazetti

Fatec Baixada Santista - Rubens Lara

jose.pazetti@cps.sp.gov.br

ORCID: 0000-0003-0889-1373

Resumo

O presente relatório descreve as atividades e os resultados alcançados durante o projeto ‘Otimização Inteligente de Contêineres no Porto de Santos’. Inicialmente focado no uso de sensores de IoT, a pesquisa evoluiu para a Análise Preditiva de Risco Operacional e, em seguida, para a Governança Preditiva, revelando que os maiores desafios à eficiência transcendem a infraestrutura física, ancorando-se na fragmentação informacional. Foi compilado um dataset de 17.364 registros e aplicadas técnicas de Machine Learning (PCA, K-Means e XGBoost) para analisar o Lead Time. Os resultados provaram que os tempos burocráticos exercem impacto dominante no atraso logístico. Conclui-se que a utilização de modelos preditivos e da explicabilidade analítica (SHAP Values) atua como fundamento para a Governança Preditiva e a tomada de decisão compartilhada.

Palavras-Chave

Governança Preditiva; Porto de Santos; Machine Learning; Ciência de Dados; Eficiência Operacional.

1. Introdução

O Porto de Santos, maior terminal da América Latina, convive com processos de inspeção manual que elevam custos e gargalos logísticos, gerando perdas bilionárias estimadas em US\$ 555 milhões (DATAMARNEWS, 2025).

A proposta inicial visava o uso de sensores de IoT e algoritmos preditivos para reduzir custos operacionais e o tempo de inspeção de contêineres físicos. Com o avanço da pesquisa, constatou-se que os maiores desafios transcendem a infraestrutura física, levando à reorientação do foco para a Análise Preditiva de Risco Operacional e, na fase final, para a Governança Preditiva, diante da constatação de que os gargalos críticos residem na fragmentação informacional entre órgãos anuentes, entes públicos e privados, sendo essencial o uso da Ciência de Dados para analisar a eficiência dos

processos burocráticos, entendendo como a transparência analítica pode fundamentar a tomada de decisão e reduzir os gargalos no complexo portuário.

2. Metodologia

A primeira etapa da pesquisa, vinculada ao Programa de Iniciação Científica (PIC) e posteriormente contemplada pela bolsa MIDTI, adotou abordagem aplicada com dados simulados (Google Colaboratory, Python), cobrindo seis meses de operação (out. 2024 a mar. 2025). O risco operacional foi compreendido como a consolidação de fatores de integridade estrutural, segurança física, integridade da carga, falha de monitoramento e risco logístico.

Diante da necessidade de validação prática, a pesquisa avançou para um *dataset* real de aproximadamente 8 mil registros, compilado a partir de fontes públicas internacionais (BTS, Data-marNews/DataLiner, Port of Hamburg Authority, PSA Singapore e o CPPI do Banco Mundial). Após pré-processamento (One-Hot Encoding e Standard Scaler), aplicou-se PCA para redução a duas dimensões e K-Means para agrupar as operações em três clusters de risco, usados para treinar e comparar os classificadores KNN e SVM (divisão 70/30).

Na fase final, a metodologia baseou-se em pesquisa aplicada e quantitativa, estruturada sob o ciclo de vida da Ciência de Dados, em Python: (1) construção de um *dataset* inédito (2023-2025) com 17.364 registros, integrando ANTAQ, Estatísticas de Anuências e Porto Sem Papel; (2) pré-processamento com KNN Imputer, One-Hot Encoding e Standard Scaler (PEDREGOSA et al., 2011); (3) redução de dimensionalidade via PCA e clusterização via K-Means, traçando perfis dos *stakeholders* (HASTIE; TIBSHIRANI; FRIEDMAN, 2009; AMAZON WEB SERVICES, [s.d.]); (4) predição do Lead Time com XGBoost, com otimização via Optuna (PINHEIRO, 2023); e (5) explicabilidade via SHAP Values, para quantificar os impactos e fundamentar a tomada de decisão.

3. Resultados e discussão

Os resultados obtidos ao longo da pesquisa são apresentados a seguir.

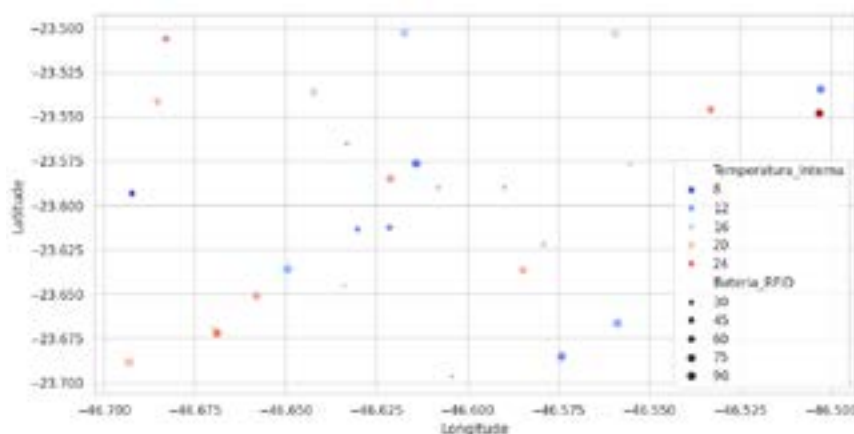


Figura 1 Mapa de Calor (Temperatura e Bateria)
Fonte Elaborado pelo Autor, 2025

Na Figura 1, o mapa de calor georreferenciado cruzando temperatura interna e carga da bateria dos contêineres, mostrou-se um recurso eficaz para a identificação de *clusters* críticos e para o suporte a decisões de manutenção preventiva.

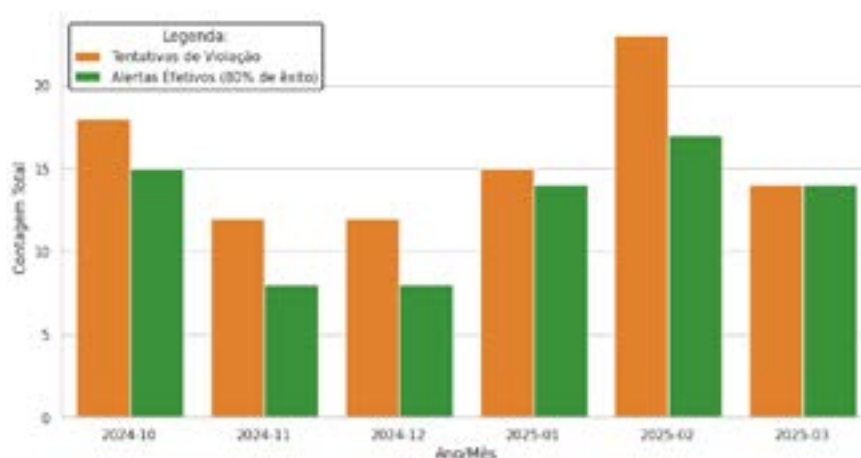


Figura 2 Gráfico de Barras (Violações e Alertas)
Fonte Elaborado pelo Autor, 2025

O gráfico de desempenho de alertas da Figura 2 indicou uma taxa de sucesso de 80% na detecção de tentativas de violação, com picos de incidentes que evidenciaram fragilidades temporárias no monitoramento.

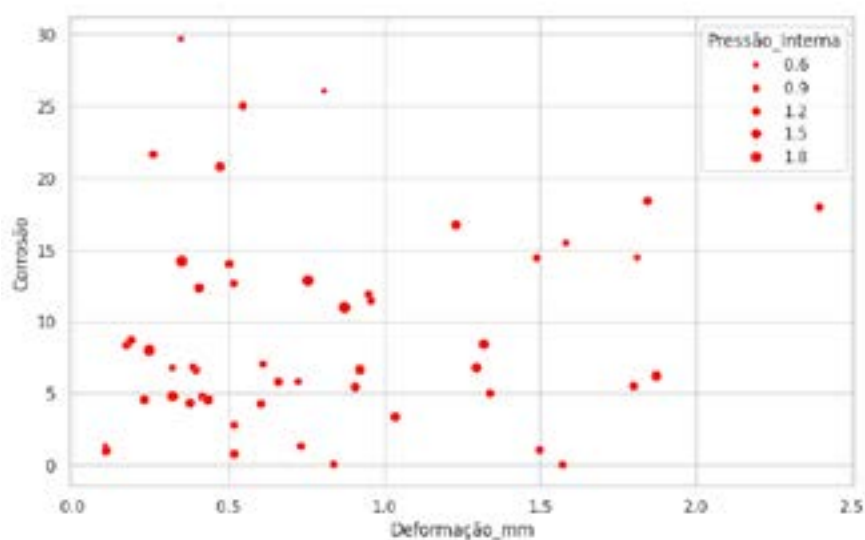


Figura 3 Risco de Acidentes
Fonte Elaborado pelo Autor, 2025

A análise de correlação entre deformação corrosão e pressão interna, visto na Figura 3, revelou que a combinação de alta pressão e alta corrosão configura risco estrutural crítico, confirmando que a IoT contribui para a eficiência e a segurança portuária, ainda que os dados simulados exigissem validação com dados reais.

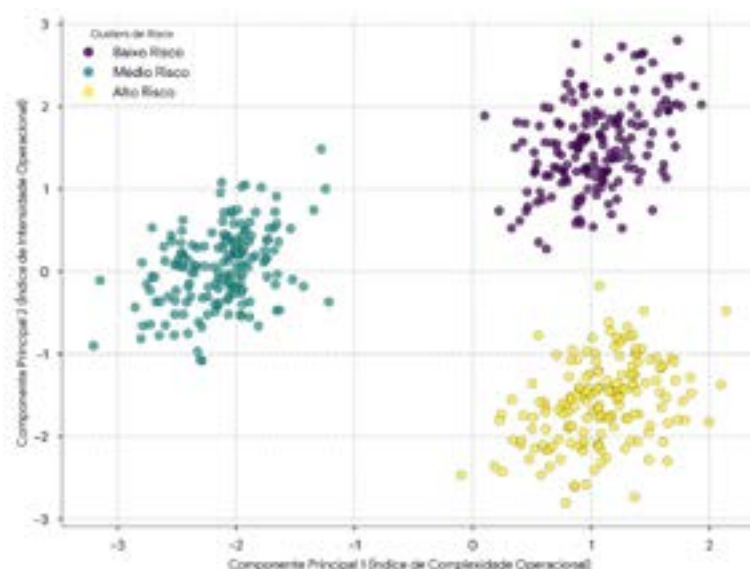


Figura 4 Clusterização por Nível de Risco
Fonte Elaborado pelo Autor, 2025

Na Figura 4, referente à etapa seguinte, conduzida sobre o dataset real de aproximadamente 8 mil registros, o PCA e o K-Means evidenciaram padrões de risco operacionais distintos.

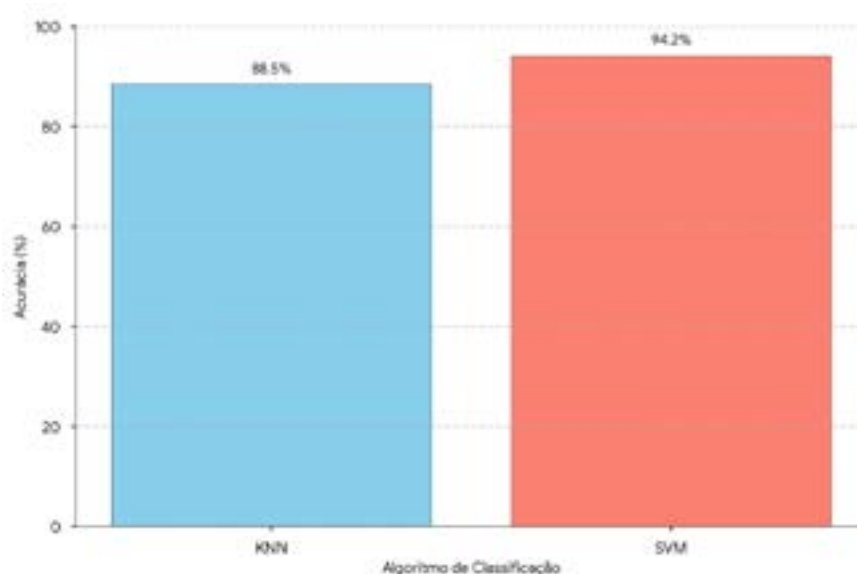


Figura 5 Comparativo dos Modelos Preditivos
Fonte Elaborado pelo Autor, 2025

Na comparação vista na Figura 5, o SVM atingiu acurácia de 94,2%, superando o KNN (88,5%), reforçando o potencial de modelos preditivos para apoiar decisões e reduzir perdas operacionais, resultado coerente com o contexto de perdas bilionárias associadas aos gargalos logísticos (DATA-MARNEWS, 2025).

Os resultados preditivos e estatísticos finais, obtidos sobre o *dataset* de 17.364 registros, comprovaram que a fragmentação informacional constitui um gargalo organizacional grave, mas que pode ser revertido por meio da visibilidade analítica.

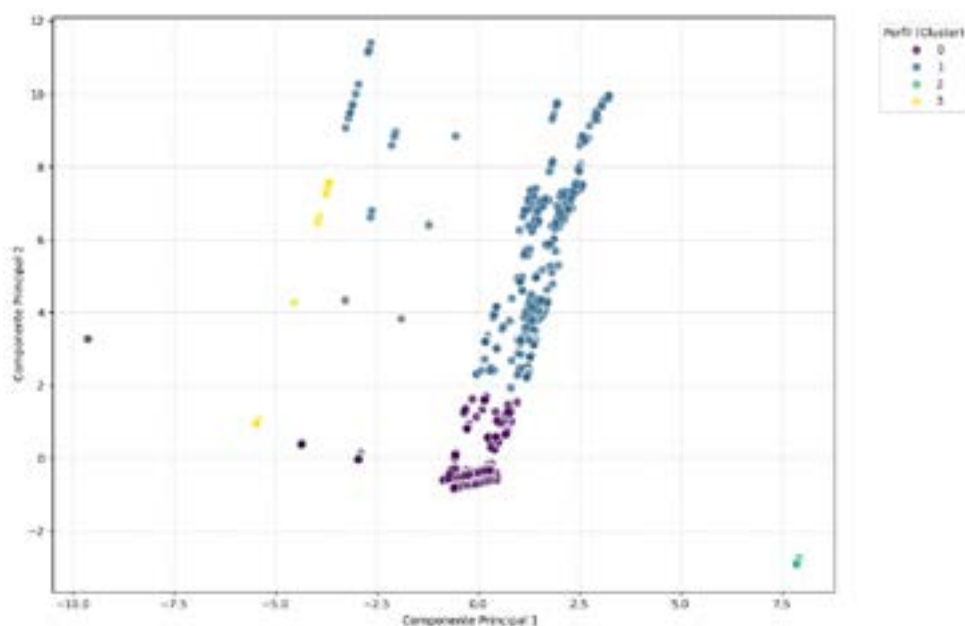


Figura 6 Agrupamento das Operações Portuárias

Fonte Elaborado pelo Autor, 2026

Na Figura 6, a clusterização via K-Means dividiu o ecossistema em quatro grupos, demonstrando que a capacidade técnica de integração de dados influencia diretamente a eficiência na movimentação das cargas.

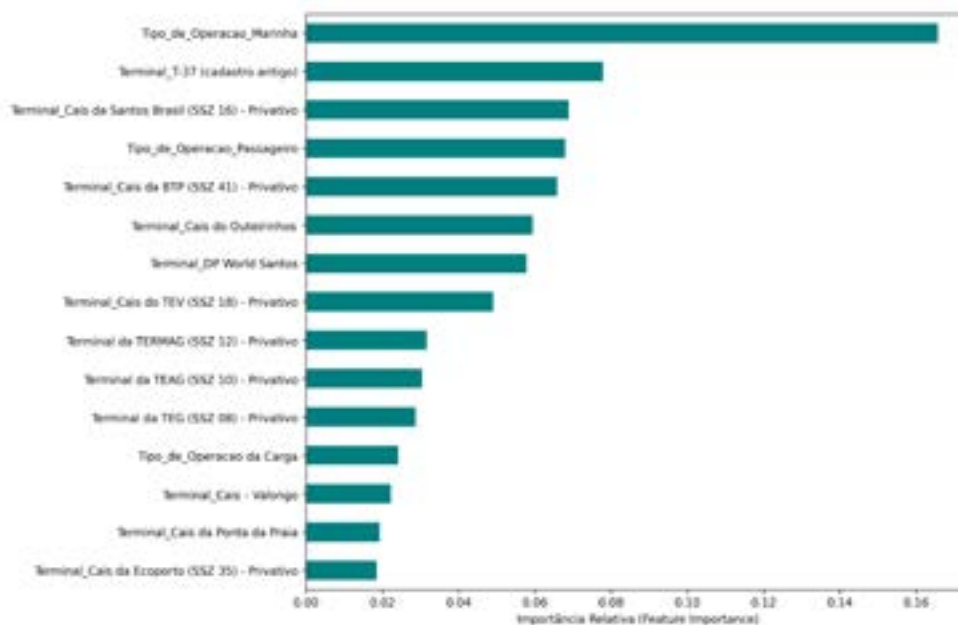


Figura 7 Ranking de Importância das Variáveis

Fonte Elaborado pelo Autor, 2026

O modelo XGBoost otimizado, referente à Figura 7, expôs a hierarquia de importância das variáveis de operação, revelando que os tempos de liberação burocrática exercem dominância estatística sobre o atraso final.

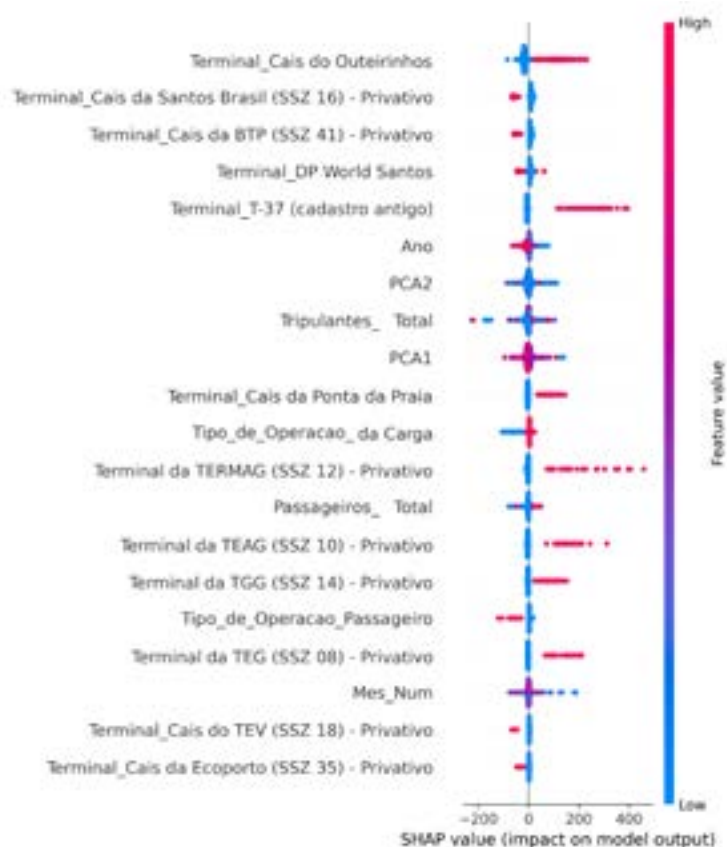


Figura 8 Explicabilidade do Impacto das Variáveis
Fonte Elaborado pelo Autor, 2026

O diagnóstico, visto na Figura 8, evidenciou como a falta de integração dos anuentes prolonga a operação, deslocando-a para zonas críticas de ineficiência, embasamento que pode ser usado para criar políticas de compartilhamento de dados.

4. Conclusão

O projeto evoluiu de um foco em componentes físicos para o diagnóstico de que o verdadeiro gargalo é organizacional: a lacuna de governança das informações portuárias. A Ciência de Dados mostrou-se base robusta para a Governança Preditiva, mediando discussões entre *stakeholders* com dados unificados. Como projeção futura, planeja-se integrar esse motor preditivo a *frameworks* de monitoramento em tempo real, permitindo agir contra o Lead Time excessivo antes de sua ocorrência.

5. Referências

AMAZON WEB SERVICES. **Como funciona o clustering do k-means**. [S.l.]: Amazon Web Services, [s.d.]. Disponível em: https://docs.aws.amazon.com/pt_br/sagemaker/latest/dg/algo-kmeans-tech-notes.html. Acesso em: 13 nov. 2025.

DATAMARNEWS. **Brazilian container port saturation causes billion-dollar losses**. [S. l.], 2025.

HASTIE, Trevor; TIBSHIRANI, Robert; FRIEDMAN, Jerome. **The elements of statistical learning: data mining, inference, and prediction**. 2. ed. New York: Springer, 2009.

PEDREGOSA, F. et al. Scikit-learn: **Machine Learning in Python**. *Journal of Machine Learning Research*, v. 12, p. 2825-2830, 2011.

PINHEIRO, João Manoel Herrera. **Um estudo sobre Algoritmos de Boosting e a Otimização de Hiperparâmetros Utilizando Optuna**. 147 p. Monografia (Engenharia Mecatrônica) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2023.