

**COMO A MINERAÇÃO DE DADOS PODE SER UTILIZADA PARA IDENTIFICAR
PADRÕES CRIMINAIS ESPECÍFICOS NA CIDADE DO NATAL/RN**
Com base na teoria da escola de Chicago sobre a criminologia

Efrain Pantaléon Matamoros¹
Horácio Betcel Guimarães²
Orivaldo Vieira de Santana Júnior³
Stefani Leite Cavalcanti⁴
Wagner Márcio Marques Cabral⁵

RESUMO:

A ocorrência constante de crimes no Brasil e no Estado do Rio Grande do Norte expõe a fragilidade do sistema de segurança pública e os novos desafios a serem enfrentados, ensejando o uso de técnicas avançadas para buscar e retirar *insights* de grande valor para serem utilizadas aos gestores no momento de definir políticas públicas que tenham relação com o setor de segurança pública. Este artigo sugere uma perspectiva moderna que une mineração de dados e algoritmos de *machine learning* e o CRISP-DM. Um arranjo que busca identificar padrões e perfis dos diversos crimes denunciados na Cidade do Natal/RN e aprimorar a compreensão de suas ocorrências, por região e bairro dentro da cidade. Os números indicam que essa abordagem não é eficaz apenas como ferramenta para aperfeiçoar estratégias, mas também tem a capacidade de identificar padrões comportamentais específicos, na busca da prevenção e diminuição dos crimes. Esse enfoque aponta para um remédio eficaz para tratar a criminalidade na cidade de Natal, cidade objeto de estudo.

Palavras-chave: criminalidade, segurança pública, técnicas avançadas, mineração de dados, algoritmos de *machine learning*, CRISP-DM, padrões, perfis, Cidade do Natal RN.

¹ Doutor em Engenharia Mecânica pela Escola Politécnica da Universidade de São Paulo (EPUSP). Professor Associado da Escola de Ciências e Tecnologia da Universidade Federal de Rio Grande do Norte na área de Negócios Tecnológicos. E-mail: efrain.pantaleon@ufrn.br.

² Graduado em Administração - Faculdades Integradas do Tapajós (1992). Mestrando do Programa de Pós-graduação de Ciência, Inovação e Tecnologia da UFRN (PPgCTI). E-mail: horaciobetcel@gmail.com.

³ Doutor e mestre pelo Centro de Informática da UFPE. Professor Adjunto da ECT/UFRN e Professor do Programa de Pós-graduação em Ciência, Tecnologia e Inovação (PPgCTI). E-mail: orivaldo@gmail.com.

⁴ Graduada em Direito pela UNI-RN, especialista em Direito Penal e Processo Penal. Mestranda do Programa de Pós-graduação de Ciência, Inovação e Tecnologia da UFRN (PPgCTI). Servidora pública do ITEP/RN. E-mail: stefanicavalcanti@hotmail.com

⁵ Graduado em Direito pela Universidade Potiguar -UNP. Mestrando do Programa de Pós-graduação de Ciência, Inovação e Tecnologia da UFRN (PPgCTI). E-mail: wagner_marcio@hotmail.com.

INTRODUÇÃO

Examinar o volume crescente de dados atualmente coletados e disponibilizados representa uma oportunidade e um desafio significativo nesta era. A estatística descritiva é uma ferramenta essencial para resumir, analisar e apresentar dados de forma prática e eficaz, além disso, o conceito de Big Data envolve a coleta e análise de grandes volumes de dados confiáveis e variados, com a velocidade necessária para gerar valor adequado às soluções (Oliveira & Rangel, 2021).

Conforme os desafios se exacerbam, faz-se imperativo o desenvolvimento de novas e avançadas ferramentas para análise e extração de informações relevantes desses dados. A coleta e investigação de informações é essencial para obter *insights* relevantes e decidir em qualquer área de trabalho (Bonvini & Abreu, 2010).

Várias bases de informação fornecem dados de formas diferentes, aumentando o hermetismo das análises, neste ambiente, o processo da Mineração de Dados desponta essencialmente para compreender os diversos cenários advindos de uma vultuosidade de informações (Vale, 2023). A principal adversidade é identificar os padrões cíclicos no banco de dados (Maurício, 2023). Para tratar deste assunto, aplica-se a ferramenta de mineração de padrões, aceita como uma visão eficiente para apontar importantes *insights* (Curvelo, 2020), com a intenção é abduzir informações relevantes, para um entendimento mais apurado dos dados coletados e facilitar na escolha de decisões (Ramos & Santos, 2003).

Neste contexto, se faz necessária a constante melhoria do tratamento da análise, assim como a busca de novas técnicas e ferramentas para trabalhar com esse vasto volume de dados fornecidos.

Os algoritmos de *machine learning*, juntamente com a mineração de dados, despontam como uma resposta necessária para explorar, compreender e extrair informações desses dados. O conjunto desses procedimentos são ferramentas que possibilitam as análises além de possibilitar a criação eficaz de modelos preditivos a partir de informações provenientes de diversas fontes. A sinergia entre os algoritmos de aprendizado de máquina e a mineração de dados potencializa as análises estatísticas, oferece suporte à tomada de decisões e gera *insights* valiosos para impulsionar a eficácia das estratégias na segurança pública e privada, contribuindo para uma tomada de decisão mais precisa (Fernandes, 2019).

Em um cenário criminalidade evidenciada, o tratamento e análise dos dados disponíveis, estabelece a diferença entre o sucesso e a ineficácia das políticas

públicas escolhidas para atender os desafios diários enfrentados pela população.

PROBLEMA DE PESQUISA E OBJETIVO

Através de dados referentes aos boletins de ocorrência registrados na cidade de Natal, dispondo sobre os tipos penais; ano e mês do fato; bairro em que fora cometido e a respectiva quantidade, fornecidos pela Coordenadoria de Informações Estatísticas e Análise Criminal, órgão vinculado à Secretaria da Segurança Pública e da Defesa Social, foi possível especificar os dez crimes mais cometidos na cidade nos anos de 2022 e 2023, conforme tabela que segue:

2022

Natureza do crime	Quantidade de BO
FURTO	11.757
ESTELIONATO	11.754
ROUBO	8.566
AMEAÇA	4.042
LESÃO CORPORAL	1.616
INJÚRIA	1.171
DANO	1.144
DIFAMAÇÃO	790
DROGAS - TRÁFICO	632
CALÚNIA	604

(Tabela criada pelos autores)

2023

Natureza do crime	Quantidade de BO
FURTO	12.018
ESTELIONATO	10.885
ROUBO	8.298
AMEAÇA	4.731
INJÚRIA	2.003
LESÃO CORPORAL	1.919
DANO	1.412
DIFAMAÇÃO	883
CALÚNIA	660
VIAS DE FATO	652

(Tabela criada pelos autores)

Assim, pôde-se observar um leve aumento com relação ao registro de alguns tipos penais ou mesma a constância de alguns deles.

Esta repetição de ocorrências levanta questões sobre a eficácia das políticas de segurança pública adotadas e sobre possíveis falhas na prevenção e combate desses crimes.

Assim, surge a necessidade de compreender quais fatores estão contribuindo para essa constância nos registros e como as autoridades podem intervir de maneira mais eficiente.

Assim, o presente artigo propõe a aplicação de algoritmos de aprendizado de máquina em conjunto com o processo de mineração de dados, com o objetivo de identificar padrões e perfis dos diversos crimes na Cidade do Natal RN, aprimorar a compreensão de suas ocorrências por região e bairro dentro da cidade e mostrar que a inteligência artificial pode ajudar no combate ao crime subsidiando decisões sobre políticas públicas na área da segurança pública.

FUNDAMENTAÇÃO TEÓRICA

O artigo tem uma de suas bases à ciência da criminologia conceituada como “uma ciência autônoma, empírica, indutiva, interdisciplinar, zetética, que tem como objetos de estudo o delito, o delinquente, a vítima e o controle social, com a finalidade de, a partir da experiência e observação, explicar, compreender e mudar a realidade criminal, com a função de subsidiar as tomadas de decisões e demais ciências enxergando o crime (ou fenômeno criminal) como um fato social” (ANDRADE, 2023, p.17).

Sabe-se que a principal função da criminologia é entender o crime como fenômeno social, com intuito de subsidiar a tomada de decisões dentro da sociedade, em especial pelos poderes públicos sobre a política criminal.

A criminologia tem uma interpretação diferente, vendo-o como “um fenômeno social, comunitário e que se mostra como um “problema” maior, a exigir do pesquisador uma empatia para se aproximar dele e o entender em suas múltiplas facetas”. Possuindo, assim, os seguintes elementos: incidência massiva na sociedade; incidência aflitiva, pois causa temor; persistência temporal; consciência social do que deve ser criminalizado e a consciência geral de sua negatividade (Andrade, 2023).

A Escola de Chicago desempenhou um papel fundamental no

desenvolvimento da criminologia moderna, influenciando tanto a teoria quanto a pesquisa empírica. Originada no início do século XX, na Universidade de Chicago, a Escola de Chicago se destacou por sua abordagem inovadora e interdisciplinar, que combinava sociologia, criminologia e estudos urbanos para entender os fenômenos criminais.

A teoria da desorganização social é uma das contribuições mais importantes da Escola de Chicago para a criminologia. Esta teoria sugere que a criminalidade é um produto da deterioração dos valores sociais e das instituições comunitárias em áreas urbanas Bryant (2014) explica que a desorganização social ocorre quando a capacidade de uma comunidade de autorregular-se e manter normas sociais é enfraquecida, levando ao aumento da delinquência e do crime.

Em resumo, a Escola de Chicago trouxe contribuições significativas para a criminologia através de suas abordagens teóricas e metodológicas inovadoras, que continuam a influenciar estudos contemporâneos sobre crime e sociedade. Suas teorias sobre desorganização social, análise espacial e metodologias empíricas permanecem relevantes e são frequentemente utilizadas para entender a dinâmica do crime em contextos urbanos.

A cidade de Natal, capital do Rio Grande do Norte, possui uma população de cerca de 1.556.413 habitantes em 2024, com alta densidade demográfica nas áreas centrais e expansão nas periferias. Bairros como Tirol, Petrópolis, Alecrim e Ponta Negra concentram grande parte da população, sendo conhecidos pela boa infraestrutura e oferta de serviços, com destaque para Alecrim como centro comercial e Ponta Negra pelo turismo (SEMURB).

Nas áreas periféricas, como Felipe Camarão e Nossa Senhora da Apresentação, o crescimento populacional é acelerado, mas enfrenta desafios significativos relacionados à infraestrutura urbana e serviços públicos, como transporte inadequado e menor acesso a serviços essenciais. Essas regiões refletem as características comuns de periferias em capitais brasileiras, com uma população crescente e necessidades urbanas insatisfeitas (IBGE, 2022).

A distribuição de renda em Natal é desigual, com bairros centrais como Tirol e Petrópolis apresentando renda mais alta e melhor acesso a empregos formais, enquanto bairros periféricos como Felipe Camarão e Nossa Senhora da Apresentação enfrentam maior vulnerabilidade econômica e social. Zonas como Ponta Negra, influenciada pela atividade turística, mostram variação de renda, e os dados indicam

a necessidade de políticas públicas que reduzam essas disparidades socioeconômicas (Dialoga RN, 2021).

METODOLOGIA

A mineração de dados é o processo de descoberta de informações acionáveis em grandes conjuntos de dados, utilizando análises matemáticas para identificar padrões e tendências. Esses padrões geralmente não podem ser descobertos por meio de métodos tradicionais de exploração de dados, devido à complexidade das relações ou ao grande volume de informações. Trata-se de uma técnica avançada que permite a identificação de padrões subjacentes, fornecendo suporte valioso para a tomada de decisões (Silva, 2020).

O CRISP-DM (Cross-Industry Standard Process for Data Mining) é uma metodologia amplamente reconhecida e utilizada para a condução de projetos de mineração de dados. Este estudo aplicou o CRISP-DM para analisar a criminalidade em Natal/RN, seguindo suas seis fases principais: compreensão do negócio, compreensão dos dados, preparação dos dados, modelagem, avaliação e implementação.

Compreensão do Negócio:

- **Objetivo:** Identificar os principais padrões e perfis criminais na cidade de Natal/RN, com o intuito de melhorar a prevenção e a resposta às ocorrências criminais.
- **Contexto:** Reunir informações sobre a estrutura atual da segurança pública, as áreas de maior incidência de crimes e os desafios enfrentados pelas autoridades locais.

Compreensão dos Dados:

- **Coleta de Dados:** Dados dos boletins de ocorrência dos crimes praticados em Natal/RN foram colhidas por fontes oficiais.
- **Exploração Inicial:** Análise exploratória dos dados para entender a qualidade e a distribuição das informações disponíveis, identificando possíveis problemas de inconsistência e lacunas nos dados.

Preparação dos Dados:

- **Limpeza dos Dados:** Tratamento de dados faltantes, remoção de duplicatas e correção de inconsistências.
- **Transformação dos Dados:** Criação de novas variáveis, normalização de

dados e segmentação das informações por ano, bairro e tipo de crime.

- **Integração dos Dados:** Combinação de múltiplas fontes de dados para formar um conjunto de dados consolidado e coerente para a análise.

Modelagem:

- **Seleção de Algoritmos:** Escolha de algoritmos de *machine learning* apropriados para a identificação de padrões e tendências, como árvores de decisão, redes neurais e *clustering*.
- **Construção dos Modelos:** Treinamento dos modelos utilizando o conjunto de dados preparado, ajustando parâmetros e validando a precisão dos resultados obtidos.
- **Teste dos Modelos:** Avaliação dos modelos em subconjuntos de dados de teste para garantir a robustez e a generalização dos padrões identificados.

Avaliação:

- **Análise dos Resultados:** Comparação dos resultados dos modelos com os objetivos iniciais do projeto, verificando se os padrões identificados são relevantes e acionáveis.
- **Feedback dos Stakeholders:** Discussão dos resultados com especialistas em segurança pública para validar a aplicabilidade e a utilidade dos insights obtidos.

Implementação:

- **Aplicação dos Modelos:** Integração dos modelos preditivos nos sistemas de suporte à decisão das autoridades de segurança pública.
- **Monitoramento e Manutenção:** Monitoramento contínuo do desempenho dos modelos e ajustes periódicos para manter a precisão e a relevância dos padrões identificados.

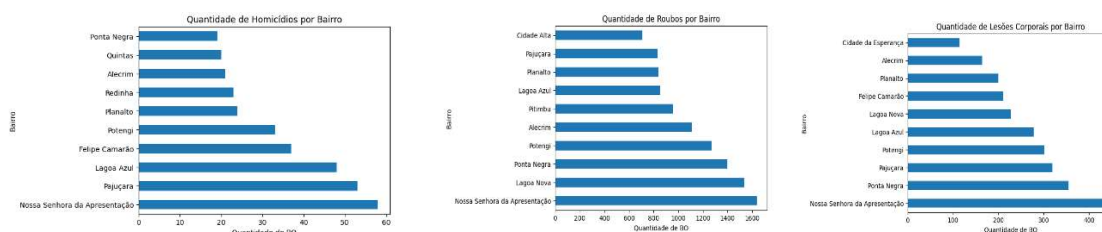
Conclusão da Metodologia: A aplicação do CRISP-DM permitiu uma abordagem estruturada e eficiente para a análise da criminalidade em Natal/RN. Através de cada uma das fases, foi possível transformar dados brutos em *insights* valiosos que podem orientar ações estratégicas na prevenção e combate ao crime. A metodologia garantiu que os modelos desenvolvidos fossem robustos, precisos e úteis para os *stakeholders* envolvidos, contribuindo para a melhoria da segurança pública na região.

Registra-se ainda que para a extração dos dados foi utilizado o *Google Colab*, ou *Colaboratory*, que permitiu a escrita e execução do código *Python* no próprio navegador. Ele forneceu todas as tabelas e gráficos apresentados no artigo.

RESULTADOS E DISCUSSÕES

Analisando as tendências dos dados apresentados é possível constatar que a quantidade de boletins de ocorrências por bairro, de 2022 a maio de 2024 se concentrou na Zona Norte de Natal.

Especificando a análise dos dados de 2022 a maio de 2024, constata-se a ocorrência de três crimes graves, sendo dois contra a integridade física e um contra o patrimônio, quais sejam: Homicídio, que tipifica a conduta de matar alguém, previsto no art. 121, do Código Penal; Lesão corporal, tipificado na conduta de ofender a integridade corporal ou a saúde de outrem, previsto no art. 129, do Código Penal e o crime de Roubo, ato de subtrair coisa alheia móvel mediante grave ameaça ou violência a pessoa, previsto no art. 157, do CP, foi possível obter os seguintes gráficos:



(Gráficos criado pelos autores)

Constata-se que dos dez bairros em que mais ocorrem os crimes de homicídio, lesão corporal e roubo, quatro ficam na Zona Norte (Nossa Senhora da Apresentação, Pajuçara, Lagoa Azul e Potengi), bairros que também representam o maior número de crimes, de forma geral, cometidos na cidade.

Não por coincidência boa parte desses bairros são apontados no artigo “Mapa da vulnerabilidade social do município de Natal RN em nível de setor censitário” como os que possuem a menor qualidade de vida, concluindo que a espacialização das áreas mostrou a existência de segregação socioespacial e que populações socialmente vulneráveis se localizam em áreas ambientalmente vulneráveis (Barbosa, 2019).

A tendência, evidenciada pelos dados analisados, é a ocorrência elevada de crimes graves em bairros com pouca infraestrutura e baixa concentração de renda.

CONCLUSÃO

A aplicação de técnicas de mineração de dados na análise de criminalidade em Natal/RN revelou-se uma abordagem poderosa e eficaz para compreender e enfrentar os desafios da segurança pública na região. Utilizando algoritmos avançados e metodologias como o CRISP-DM, foi possível identificar um padrão nas ocorrências criminais que se alinham a teoria da criminologia desenvolvida pela Escola de Chicago.

Restou constatado que em bairros com menor presença do Estado, através de infraestrutura básica, ocorrem mais crimes graves, especificados, nesse artigo, nos crimes de homicídio, lesão corporal e roubo. Todos os crimes citados envolvem violência e agressão à pessoa.

A segmentação das ocorrências por região e bairro permitiu uma compreensão mais granular das dinâmicas criminais locais, facilitando a formulação de estratégias de prevenção mais precisas e eficientes.

Os resultados deste estudo demonstram que a mineração de dados não apenas aperfeiçoa a tomada de decisões no setor de segurança pública, mas também proporciona insights valiosos sobre comportamentos e tendências criminais. Com a implementação contínua dessas técnicas, espera-se uma melhoria significativa nas ações de prevenção e combate à criminalidade, contribuindo para a redução dos índices de crimes em Natal/RN. A integração dessas tecnologias na gestão de segurança pública representa um avanço significativo, indicando um caminho promissor para outras regiões que enfrentam desafios semelhantes.

REFERÊNCIAS

1. ANDRADE, A.; MEDEIROS, D. ***Criminologia decifrada***. Coordenação Cláudia Barros, Filipe Ávila, Rogério Greco. 2. ed. Rio de Janeiro: Método, 2023.
2. AZEVEDO, A.; SANTOS, M. F. **KDD, SEMMA and CRISP-DM: A Parallel Overview**. In: ***Proceedings of the IADIS European Conference on Data Mining 2008***, p. 182-185, 2008.
3. BARBOSA, I. R.; GONÇALVES, R. C. B.; SANTANA, R. L. **Mapa da vulnerabilidade social do município de Natal RN em nível de setor censitário**. 2019. Disponível em: <https://www.revistas.usp.br/jhgd/article/download/157749/153497/351108>. Acesso em: 08 jul. 2024.
4. BERRY, M.; LINOFF, G. ***Data Mining Techniques for Marketing, Sales, and Customer Relationship Management***. John Wiley & Sons, 2004.

5. BREIMAN, L. Random forests. ***Machine Learning***, v. 45, n. 1, p. 5-32, 2001.
6. CASTANHEIRA, L. G. **Aplicação de Técnicas de Mineração de Dados em Problemas de Classificação de Padrões**. 2008. Dissertação (Mestrado em Engenharia Elétrica) - Universidade Federal de Minas Gerais (UFMG), Belo Horizonte.
7. CHAPMAN, P.; CLINTON, J.; KERBER, R.; KHABAZA, T.; REINARTZ, T.; SHEARER, C.; WIRTH, R. **CRISP-DM 1.0: Step-by-step data mining guide**. SPSS Inc., 2000. Disponível em: <https://www.the-modeling-agency.com/crisp-dm.pdf>. Acesso em: 07 jul. 2024.
8. CORCOVIA, L. O.; ALVES, R. dos S. **Aprendizagem de Máquina e Mineração de Dados: Avaliação de Métodos de Aprendizagem**. *Interface Tecnológica*, 2019.
9. CURVELO, R. **Data Mining: Entenda o que é e como funciona**. 2020. Disponível em: <https://br.hubspot.com/blog/marketing/data-mining>. Acesso em: 07 jul. 2024.
10. ENTENDA o CRISP-DM, suas etapas e como de fato gerar valor com essa metodologia. Disponível em: <https://www.preditiva.ai/blog/entenda-o-crisp-dm-suas-etapas-e-como-de-fato-gerar-valor-com-essa-metodologia>. Acesso em: 07 jul. 2024.
11. LIMA, A. **Mineração de Dados e Análise no comportamento de vendas de passagens fluviais na Amazônia**. Disponível em: <https://ojs.revistadelos.com/ojs/index.php/delos/article/view/1195/1030>. Acesso em: 07 jul. 2024.
12. MARISCAL, G.; MARBÁN, Ó.; FERNÁNDEZ, C. **A survey of data mining and knowledge discovery process models and methodologies**. *Knowledge Engineering Review*, v. 25, n. 2, p. 137-166, 2010. DOI: 10.1017/S0269888910000032.
13. WORLD Population Review. 2024. Disponível em: <https://worldpopulationreview.com/countries>. Acesso em: 02 jul. 2024.