

Universidade de Pernambuco
Programa de Pós-Graduação em Engenharia da
Computação (PPGEC)

Proposta de Dissertação de Mestrado

Área: Inteligência Computacional

Título: Análise e Descoberta de Padrões em Dados Abertos na Área da Saúde

Orientador(a): Dimas Cassimiro do Nascimento Filho
(dimas.cassimiro@ufape.edu.br)

Descrição:

O crescente volume de dados disponíveis publicamente representa uma fonte de dados relevante para a descoberta de padrões e análises inteligentes que podem beneficiar a sociedade em diversas áreas críticas, tais como a saúde pública. A análise de padrões em dados abertos é crucial para compreender fenômenos complexos, identificar tendências emergentes e apoiar a tomada de decisões. Este tema de pesquisa visa contribuir para o avanço da capacidade de explorar e interpretar conjuntos de dados abertos disponíveis na área da saúde, promovendo assim a utilização eficaz dessas informações para benefício da sociedade.

O objetivo deste tema consiste em investigar métodos avançados de análise de dados, técnicas de mineração de dados e algoritmos de Inteligência Computacional para identificar padrões significativos e tendências em conjuntos de dados abertos. Além disso, pretende-se desenvolver novas estratégias para visualização de dados que possam auxiliar na compreensão e interpretação desses dados. Contribuindo assim, para a disseminação de dados disponíveis em bases de dados públicas e para aprimorar o processo decisório baseado em dados abertos disponíveis.

Dentre as contribuições específicas que podem ser desenvolvidas, podem ser destacadas:

- 1) A criação de ferramentas para visualização de dados abertos;
- 2) A aplicação e adaptação de técnicas de Ciência de Dados para análise descritiva de dados abertos;
- 3) A aplicação e combinação de modelos de Inteligência Computacional para prever resultados em dados abertos;
- 4) A aplicação e adaptação de técnicas de mineração de dados para a descoberta de padrões não triviais em dados abertos;
- 5) Modelagem de bancos de dados para representação e otimização de consultas sobre dados abertos.

A identificação e compreensão de padrões não triviais em dados abertos não apenas aprimora a utilidade dos dados disponibilizados, mas também impulsiona a inovação, permitindo a formulação de políticas mais informadas e a criação de soluções personalizadas para desafios complexos. Além disso, a aplicação de técnicas de ciência de dados no contexto de dados abertos poderá produzir novas contribuições científicas, por meio da análise de uma ampla variedade de dados abertos em diferentes contextos.

Referências Bibliográficas:

- [1] NASCIMENTO, C. F. et al. Machine learning para análises preditivas em saúde: exemplo de aplicação para prever óbito em idosos de São Paulo, Brasil. *Cadernos de Saúde Pública*, v. 35, n. 7, e00050818, 2019.
- [2] JATOBÁ, A. et al. Predictive estimations of health systems resilience using machine learning. *BMC Medical Informatics and Decision Making*, v. 25, art. 267, 2025.

- [3] DHANDA, S. S. et al. Advancement in public health through machine learning: a narrative review of opportunities and ethical considerations. *Journal of Big Data*, v. 12, art. 154, 2025.
- [4] RAKERS, M. M. et al. Availability of Evidence for Predictive Machine Learning Algorithms in Primary Care: Systematic Review. *JAMA Network Open*, 2024.
- [5] SAFARIPOUR, Razieh et al. Machine Learning in Population Health: Frequent Emergency Department Utilization Pattern Identification and Prediction. 2021. Tese de Doutorado. University of Saskatchewan.
- [6] CARNEIRO, D. K. de O.; GOULART, G. S. da S.; PORTO, R. B. Big Data in the Brazilian public health sector: concept, characteristics, benefits, and challenges. *Revista ENAP*, 2022.
- [7] XAVIER, F.; BARBOSA, G. L.; MARQUES, C. C. A.; SARAIVA, A. M. Abordagem Big Data–Saúde Planetária para avaliação do Programa de Combate à Dengue. *Revista de Saúde Pública*, v. 58, n. 1, 2024.
- [8] BITTENCOURT, Cinthya; ESTIMA, Jacinto; PESTANA, Gabriel. Open data initiatives in Brazil. In: 2019 14th Iberian Conference on Information Systems and Technologies (CISTI). IEEE, 2019. p. 1-4.