

# Redes Bayesianas e Análises de Trajetórias: Uma proposta para criação de modelos indutivos penalizados

## Doutorado

Elias Carvalho, Julio César Nievola

**Contexto:** Existe uma variedade de problemas de saúde que tem gerado preocupação para a saúde pública e prejuízos para a população/governo. Conjuntos de dados clínicos são disponibilizados publicamente e utilizados para o desenvolvimento de soluções com base em técnicas de aprendizagem de máquina. A modelagem por meio de grafos e redes tem sido uma dessas soluções com destaque para as redes Bayesianas. Modelos baseados em redes Bayesianas tem limitações para identificar a intensidade das relações entre as variáveis e nem sempre são otimizados para encontrar caminhos exatos e expressar a influência de uma variável para outra. Como exemplo, após a aprendizagem da estrutura da rede, se atribui a intensidade das relações por regressões lineares simples. A análise de trajetória, tem crescido em aplicações para identificar padrões de associações hierárquicas que permitem identificar efeitos diretos e indiretos. O principal problema está na identificação desses modelos que precisam de um conhecimento teórico ou clínico sólido para garantir a adequação deste e sua aplicabilidade clínica. **Objetivo Geral:** Desenvolver um pacote no ambiente estatístico R para aliar a capacidade de geração de redes Bayesianas a partir de conjunto de dados e automatizar a criação de modelos de análise de trajetórias. **Questão de Pesquisa:** É possível criar um método computacional que permitisse aliar as duas metodologias e gerar um modelo mais robusto? Existe a possibilidade de automatizar a criação do modelo para análise de trajetórias gerando a capacidade de criação de modelos empíricos com consistência clínica baseado em mineração de dados? **Método de Pesquisa:** A metodologia deste estudo se baseia na utilização de algoritmos de aprendizagem de estrutura de Redes Bayesianas e métodos para criação de análises de trajetórias. A análise de trajetórias é um submétodo do método de modelagem de equações estruturais, um conjunto diversificado de modelos matemáticos, algoritmos e métodos estatísticos para construção de redes de dados. Esses métodos foram escolhidos por: 1) redes Bayesianas se destacam pela sua capacidade de lidar com a incerteza por seus métodos probabilísticos, 2) análise de trajetória é um método robusto por utilizar estatística multivariada, estimando efeitos diretos e indiretos das variáveis independentes, o que não ocorre na regressão linear. **Resultados Preliminares:** Foi desenvolvido o pacote “bnpa” a ser disponibilizado no

repositório do CRAN-R. A validação inicial do pacote foi realizada no desenvolvimento do modelo para o artigo "*A Spatio-Temporal Comparison of suicide mortality in Parana state, Brazil*" em processo de publicação, e por especialistas da área.

**Palavras-chave:** Análise de Trajetórias, Aprendizagem de Máquina, , BNPA, Redes Bayesianas