

APRENDIZADO DE MÁQUINA PARA SAÚDE

Inferência Causal: da correlação ao efeito

DAGs, confundimento e limites da evidência observacional

APOSTILA TÉCNICA · ENCONTRO 12

Prof. Flávio Luiz Seixas
Pós-graduação em Computação

Sumário

Sumário	2
Introdução	3
1. Observação e intervenção	3
2. Confundimento e DAGs.....	3
3. Comparabilidade em dados reais	4
4. Evidência, perturbação e sistemas que aprendem.....	4
Síntese do encontro.....	4
Referências	4

Introdução

Predizer o que costuma ocorrer após observar X não é o mesmo que estimar o que ocorreria se X fosse alterado. Essa distinção é central para perguntas sobre intervenção, tratamento e política. [1]

O encontro introduz relações causais, confundimento, grafos direcionados e estratégias de comparabilidade. Como os PDFs locais não cobrem detalhadamente todas as técnicas previstas nos slides, o texto mantém o aprofundamento compatível com a evidência disponível. [1]

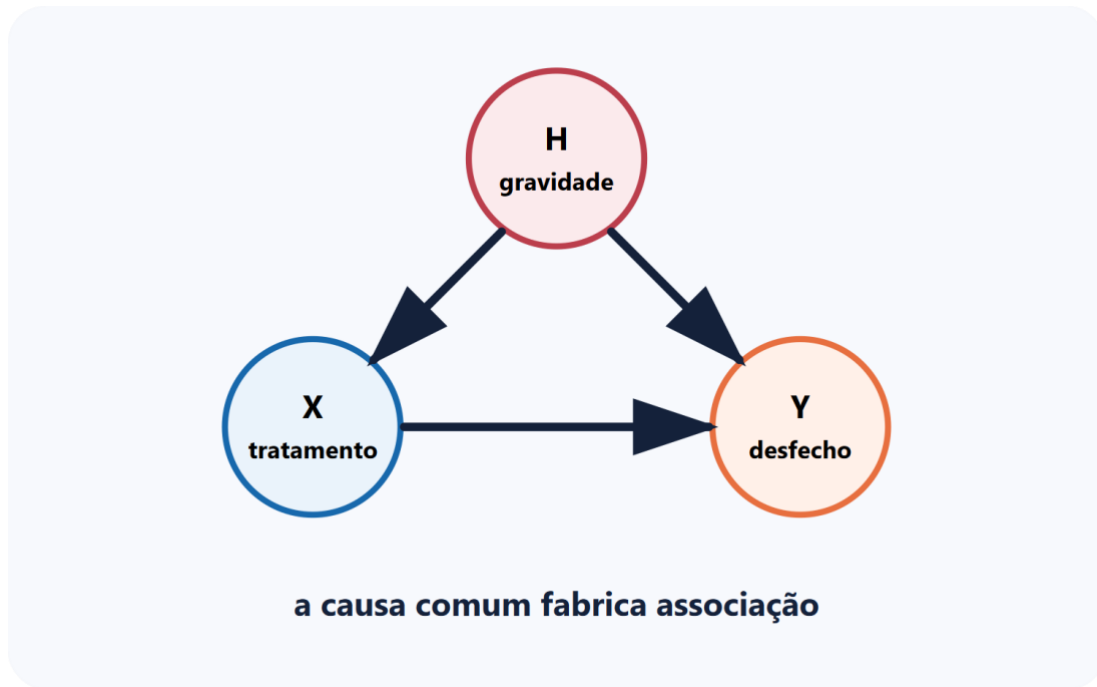


Figura 1 — Confundimento representado em DAG. Fonte: elaboração própria com base em [1] e [2].

1. Observação e intervenção

Modelos preditivos estimam variáveis depois de observar outras variáveis. Inferência causal procura relações que permitam prever consequências depois de definir ou intervir em uma variável. Diagnóstico com informações existentes é uma previsão observacional; escolher uma terapia exige uma previsão sobre consequências de ações [1].

CONCEITO-CHAVE

Modelos preditivos estimam variáveis depois de observar outras variáveis.

2. Confundimento e DAGs

Confundimento ocorre quando uma causa comum influencia exposição e desfecho, criando ou alterando uma associação. Se o fator não é medido, dados observacionais podem sustentar explicações incompatíveis. Tamanho amostral maior não remove automaticamente esse problema [2].

Grafos acíclicos direcionados representam hipóteses sobre relações causais. Eles ajudam a tornar explícitas causas comuns, mediadores e colisores e a discutir que variáveis devem ser controladas. O grafo não é descoberto apenas por correlação: depende de conhecimento e pressupostos [1].

3. Comparabilidade em dados reais

Ensaio randomizado usa alocação para criar comparabilidade esperada. Em dados observacionais, estratificação, pareamento e ponderação procuram equilibrar covariáveis medidas. O propensity score resume a probabilidade de receber a exposição dadas covariáveis observadas; ele não resolve confundidores não medidos [2].

Confundimento por indicação ocorre quando prognóstico e decisão de tratamento estão relacionados. Comparações diretas podem então atribuir ao tratamento diferenças que refletem a indicação clínica [2].

4. Evidência, perturbação e sistemas que aprendem

Intervenções ou variações externas podem oferecer informação causal que correlação isolada não fornece. Em biologia translacional, perturbar uma variável e observar mudanças auxilia a estabelecer direção entre mecanismos [3].

Um sistema de saúde que aprende combina dados do cuidado, análise, implementação e nova avaliação. O ciclo só produz evidência confiável quando pergunta, tratamento, tempo, desfecho e pressupostos causais são especificados antes da análise [4].

LIMITAÇÃO DOCUMENTAL

Emulação de ensaio-alvo, randomização mendeliana e fórmulas de propensity score não têm cobertura detalhada nos PDFs locais.

Síntese do encontro

Observação e intervenção: modelos preditivos estimam variáveis depois de observar outras variáveis.

Confundimento e DAGs: confundimento ocorre quando uma causa comum influencia exposição e desfecho, criando ou alterando uma associação.

Comparabilidade em dados reais: ensaios randomizados usam alocação para criar comparabilidade esperada.

Evidência, perturbação e sistemas que aprendem: intervenções ou variações externas podem oferecer informação causal que correlação isolada não fornece.

Referências

- [1] Visweswaran, S.; King, A. J.; Cooper, G. F. Integration of AI for Clinical Decision Support. In: Intelligent Systems in Medicine and Health. Springer, 2022.
- [2] Friedman, C. P.; Wyatt, J. C. Evaluation of Biomedical and Health Information Resources. In: Biomedical Informatics. Springer, 2021.
- [3] Butte, A. J. et al. Translational Bioinformatics. In: Biomedical Informatics. Springer, 2021.
- [4] Brixey, J. J.; Burningham, Z. Evidence-Based Informatics. In: Health Informatics. 2021.