

ALUNO: CLEITON MOYA DE ALMEIDA
EST5104 - INFERÊNCIA BAYESIANA
PROF. JOSEMAR RODRIGUES
28/03/2026

Lista 1 - Artigo Breiman

1. Qual é a principal provocação de Breiman?

A principal provocação é sobre a ênfase dada, por parte da comunidade Estatística, na modelagem dos dados usando modelos matemáticos (ou seja, a Natureza segue um modelo matemático para gerar os dados), em detrimento da abordagem algorítmica utilizada pela comunidade de Computação. Segundo o autor, esta ênfase tem levado a teorias irrelevantes, conclusões questionáveis e impedindo que o estatísticos trabalhem em uma ampla classe de problemas atuais.

2. Você concorda com a provocação? Por quê?

Concordo em partes. Eu acredito que, na época em que artigo foi escrito (2001), a provocação era bem mais relevante que hoje. Atualmente, acredito vários estatísticos já consideram e estudam também as abordagens algorítmicas. Além disso, noto que há uma maior tentativa (em ambas comunidades), de formalizar e verificar possíveis equivalências entre os modelos algorítmicos com os modelos matemáticos clássicos. Atualmente, vários estatísticos estão dedicados a estudar formalismos e garantias para os modelos de Aprendizado de Máquina. Penso então que o artigo foi bastante importante para o cenário atual.

3. Você acredita que Modelagem de Dados é uma abordagem ruim? Por quê?

Acredito que é uma ferramenta importante, mas que naturalmente possui limitações e não é uma "bala de prata" para se resolver todos os problemas. Por exemplo, quando o foco é entender as relações entre as variáveis (ou seja, necessita-se de interpretabilidade, e não somente predição), a Modelagem de Dados continua sendo importante, mesmo tendo a limitação de não conseguir modelar e responder perguntas sobre causa e efeito.

4. O paradigma Bayesiano pode ser considerado Modelagem de Dados? Por quê?

Sim. Porque ele também utiliza um modelo matemático para explicar os dados e responder perguntas sobre os mesmos através de associações (sem considerar causa e efeito). A diferença é que o paradigma bayesiano incorpora também no modelo a opinião subjetiva do analista (priori), enquanto que a abordagem clássica considera somente a informação proveniente dos dados.

5. Qual é a principal vantagem da Modelagem Algorítmica?

A principal vantagem é em aplicações com bases de dados grandes dimensões, ou dados mais complexos (que exibem relações não lineares), e sobretudo quando o objetivo é predição, e não interpretação. Nestas bases, os modelos de aprendizado de máquina geralmente apresentam melhor desempenho na tarefa de predição que os modelos clássicos.

6. O método Bayesiano não paramétrico é uma alternativa à Modelagem de Dados?

Sim, é uma alternativa à modelagem clássica (frequentista) e também à modelagem bayesiana paramétrica. Ela pode apresentar vantagem em bases de dados complexas, com relações não-lineares difíceis de serem modeladas com modelos clássicos paramétricos. Porém, também sofre da limitação de não ser capaz de responder perguntas de causa e efeito.