

## Aprendizado de Máquina

# Aula 2: Algoritmos de Aprendizado de Máquina

André C. P. L. F de Carvalho  
ICMC/USP

[andre@icmc.usp.br](mailto:andre@icmc.usp.br)



# Tópicos

- Utilizando Aprendizado de Máquina
- Viés indutivo
- Algoritmos de Aprendizado de Máquina
- Dilema viés-variância
- Atributos preditivos
- Parâmetros e híper-parâmetros

# Utilizando Aprendizado de Máquina

Modificar  
algoritmo de AM

Lidar com dados  
desbalanceados

Engenharia  
de atributos

Selecionar  
atributos



Adaptado de Rick Caruana, Research  
opportunities in AutoML Microsoft Research

Ajustar hiper-  
parâmetros

Lidar com  
valores ausentes

Verificar  
overfitting

Descobrir bugs  
no código

# AM de ponta-a-ponta

- Utilização de algoritmos de Aprendizado de Máquina em aplicações reais
  - Incluem mais que indução de modelos
  - Vai da criação de um conjunto de dados até a validação do modelo induzido, incluindo:
    - Gerenciamento, integração e pré-processamento de dados
    - Seleção e ajuste de modelos
    - Validação por especialista do domínio
    - Framework e pipeline

# Algoritmos de AM

- Induzem modelos (hipóteses) a partir de um conjunto de dados
- Dados precisam ser:
  - Estruturados
  - Representativos
  - De boa qualidade
- Algoritmos de AM indutivo possuem um viés
  - Tendência a privilegiar uma dada hipótese ou conjunto de hipóteses



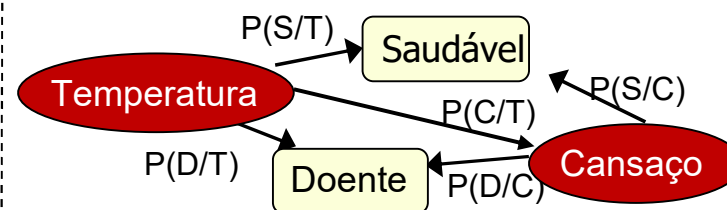
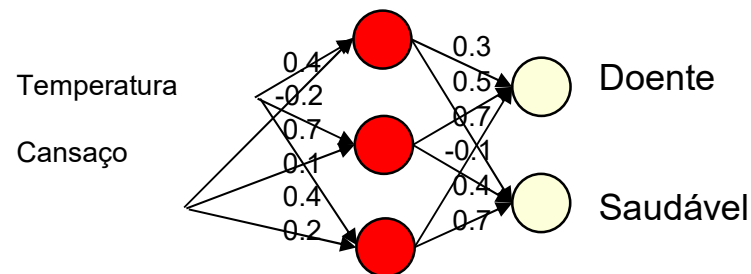
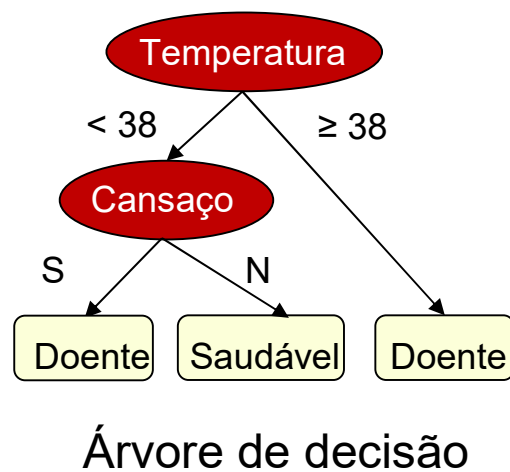
# Viés indutivo

- Algoritmos de AM precisam ter um viés indutivo
  - Necessário para restringir o espaço de busca
  - Sem viés não há generalização
    - Regras / equações seriam especializados para os dados usados para a indução modelos

# Viés indutivo

- Viés de preferência ou busca
  - Como as hipóteses são pesquisadas no espaço de hipóteses
  - Preferência de algumas hipóteses sobre outras
    - Ex.: preferência por hipóteses simples (curtas)
- Viés de representação ou linguagem
  - Define o espaço de busca ou de hipóteses
  - Restrição das hipóteses que podem ser geradas
    - Ex.: hipóteses no formato de árvores de decisão

# Viés de representação





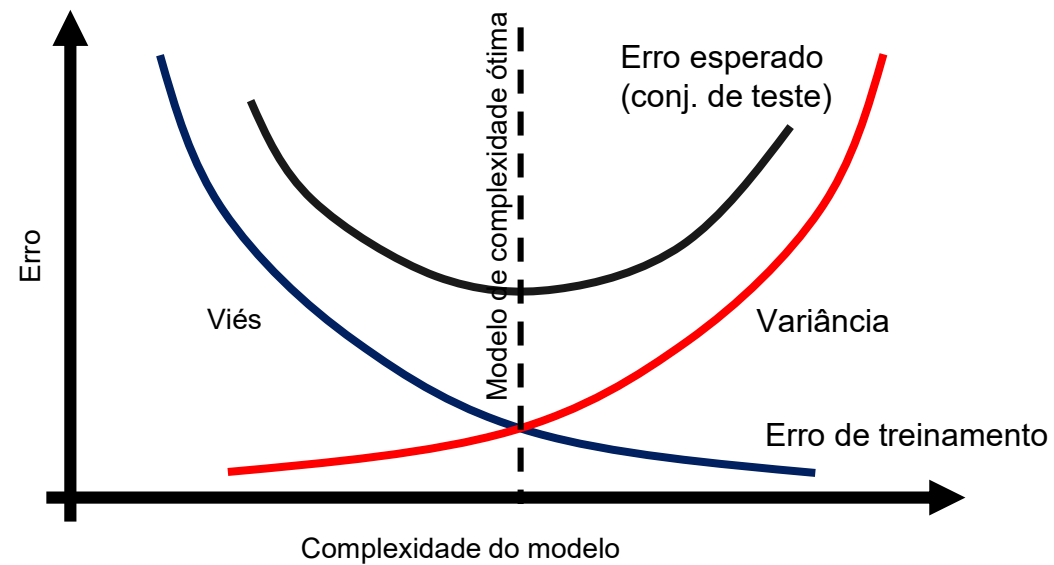
# Viés de representação

- Influência da representação no viés
  - Ex.: Hipótese é representada por uma regra que apresenta apenas conjunções
    - Regras com negações ou disjunções não fazem parte do espaço de hipóteses
    - Reduz consideravelmente espaço de possíveis hipóteses

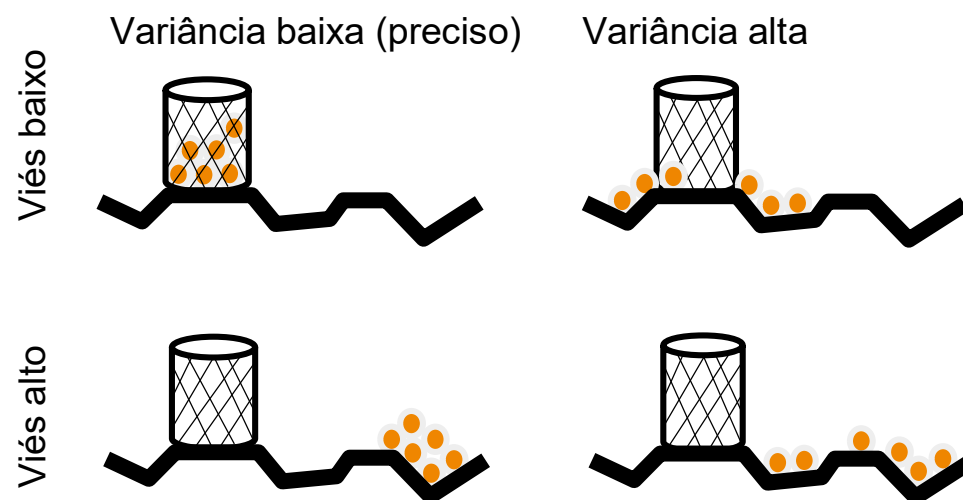
# Algoritmos de AM

- Fontes de erro de algoritmos AM
  - Viés (quanto mais alto, pior)
    - Quando algoritmo aprende um modelo incorreto
    - Associado a *underfitting*
  - Variância (quanto mais alto, pior)
    - Quando algoritmo presta atenção a detalhes sem importância
    - Associado a *overfitting*
- Precisam ser reduzidos

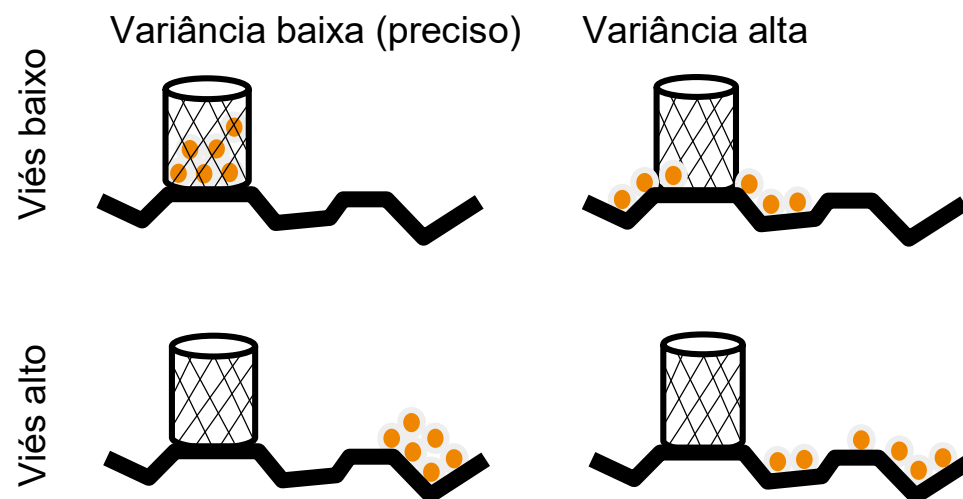
# Viés-Variância trade-off



# Viés-Variância trade-off



# Viés-Variância trade-off



Variância alta: Alucinação

Viés alto: Desatenção, não aprende

Acertar sempre no alvo, de qualquer posição: viés baixo e variância baixa

# Aprender é achar padrões

- Formar e lembrar de novos conceitos
  - E adaptar conceitos conhecidos
- Algoritmo que não aprende
  - Não consegue induzir modelo que se ajusta (fit) aos dados
    - Underfitting
  - Não presta atenção aos dados
    - Desatento



# Aprender é esquecer

- Aprender é tanto esquecer os detalhes quanto lembrar das partes importantes
  - Prestar muito atenção a detalhes leva a overfitting do modelo induzido
    - Algoritmo de AM acha um padrão nos dados que não é verdadeiro no mundo real
  - Presta atenção a detalhes
    - Alucinação

# Bom algoritmo de AM

- Está sempre percorrendo um caminho estreito entre:
  - Alucinação (*overfitting*)
  - Desatenção (*underfitting*)
- Buscando o melhor compromisso que reduz reduzir ambos



# Algoritmos preditivos

- Função de aproximação (fronteira de decisão) pode ser muito complexa
  - Difícil de obter por técnicas tradicionais
- Algoritmos de AM utilizam heurísticas para procurar essas funções
- Atributos extraídos podem não representar bem a tarefa
  - Dificultando a indução de bons modelos

# Algoritmos preditivos

- Atributos preditivos procuram descrever a tarefa a ser resolvida
  - Em geral, quanto mais atributos são extraídos, melhor
    - Facilitam indução de bons modelos
  - No entanto
    - Dificultam visualizar distribuição dos dados
    - Podem incluir atributos irrelevantes, redundantes. ...
    - Maldição da dimensionalidade

# Algoritmos preditivos

- Podem ser agrupados por diferentes critérios
  - Baseados em distâncias
    - K-NN
  - Baseados em otimização (conexionistas)
    - Redes Neurais
  - Baseados em probabilidade
    - Naive Bayes
  - Baseados em procura (lógicos)
    - Indução de árvores de decisão

# Algoritmos preditivos

- Podem ser agrupados por diferentes critérios
  - Baseados em distâncias
    - K-NN
  - Baseados em otimização
    - RNs
  - Baseados em probabilidade
    - NB
  - Baseados em procura (lógicos)
    - Indução de ADs

Geométricos



# Ajuste de (hiper-)parâmetros

- Algoritmos de AM ajustam valores de um conjunto de parâmetros
  - Parâmetros do modelo a ser gerado
  - Cada conjunto de valores pode gerar um modelo diferente
- Algoritmos de AM possuem hiper-parâmetros
  - Definem que modelos o algoritmo pode gerar
  - Ajustados por quem está aplicando o algoritmo
    - Buscando induzir modelos com melhor desempenho (preditivo)

# Ajuste de (hiper-)parâmetros

- Desempenho dos modelos induzidos depende dos valores de seus hiper-parâmetros
- Quanto mais hiper-parâmetros
  - Maior a flexibilidade para indução de modelos
  - Mais difícil o ajuste do algoritmo
- Muitos hiper-parâmetros, tudo é possível

*Com 4 parâmetros eu posso modelar um elefante, e com 5 eu posso fazê-lo mover sua tromba*



John Von Neumann

# Conclusão

- Aprendizado de Máquina
- Viés indutivo
- Algoritmos de Aprendizado de Máquina
- Dilema viés-variância
- Algoritmos preditivos
- Parâmetros e hiper-parâmetros

# *Final da Apresentação*

Copyright © 2020. Todos os direitos reservados ao CeMEAI-USP. Proibida a cópia e reprodução sem autorização

