

**Aprendizado de Máquina**

# **Aula 4: Algoritmos de Aprendizado de Máquina**

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

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



**CeMEAI**  
CEPID - Centro de Ciências  
Matemáticas Aplicadas à Indústria

# Tópicos a serem abordados

- Algoritmos de aprendizado de máquina
- Viés indutivo
- Dilema viés-variância
- Atributos preditivos
- Parâmetros e híper-parâmetros

# Algoritmos de aprendizado de máquina

- Para cada tarefa de aprendizado de máquina, podemos pensar em um grande número de aplicações
  - Cada aplicação é representada por um conjunto de dados
- Para encontrar uma boa solução para uma aplicação precisamos encontrar um modelo que se ajuste bem ao seu conjunto de dados
  - Em aplicações preditivas, deve produzir a saída desejada para cada entrada
- A forma de procurar por esses modelos é descrita por um algoritmo de aprendizado de máquina
  - Sequência de passos que definem como procurar por um bom modelo para um conjunto de dados
    - Conjunto de regras que definem como ajustar os parâmetros do modelo

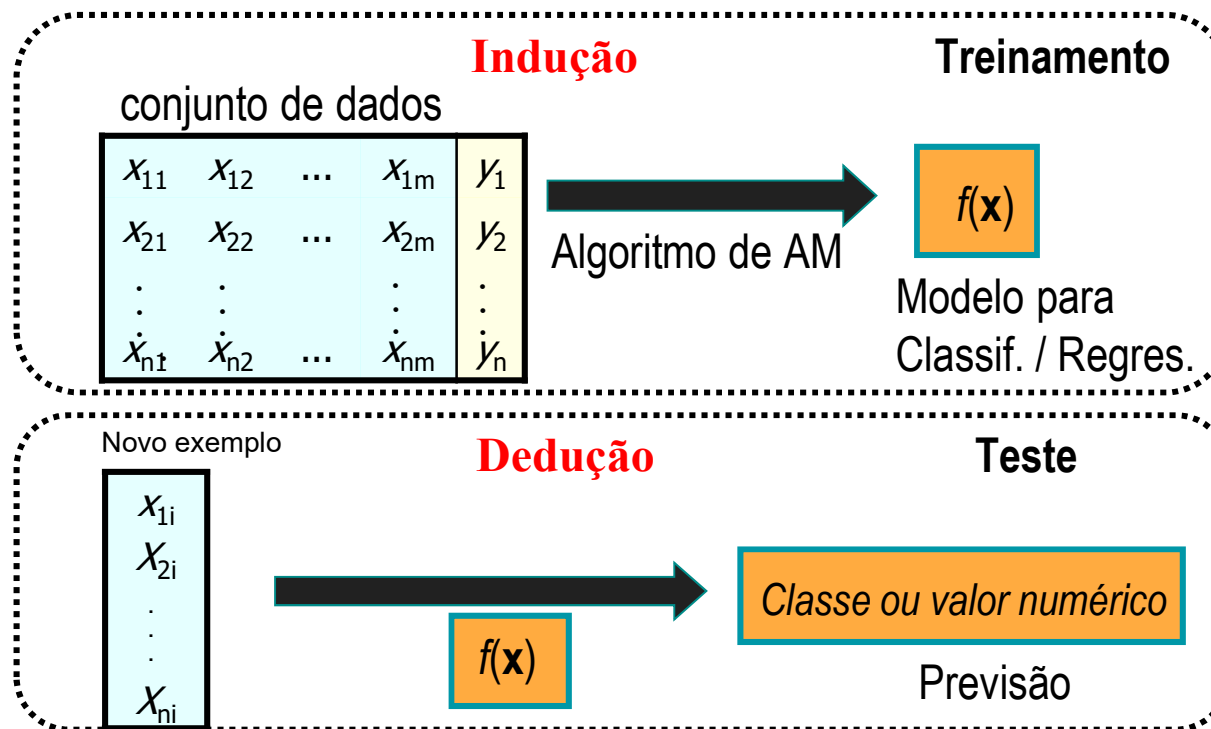
# Algoritmos de aprendizado de máquina

- Podem seguir diferentes paradigmas
  - Supervisionado
    - Podem ser usados em tarefas preditivas (mais comum) ou descritivas
    - Algoritmo de aprendizado supervisionado usa atributos preditivos e alvo
  - Não supervisionado
    - Podem ser usados em tarefas descritivas (mais comum) ou preditivas
    - Algoritmo de aprendizado não supervisionado usa apenas atributos preditivos
  - Semi-supervisionado
    - Caso especial: aprendizado ativo
  - Por reforço

# Algoritmos supervisionados

- Em geral induzem modelos (funções) preditivas
  - Fase de treinamento
  - Usando dados de treinamento rotulados
- Modelo pode ser aplicado a novos dados (predição)
  - Fase de teste
  - Prediz rótulo para cada exemplo de teste
- Principais tarefas onde são aplicados:
  - Classificação
  - Regressão

# Algoritmos supervisionados





# Algoritmos de AM

- Induzem modelos (funções, hipóteses) a partir de um conjunto de dados
  - Idealmente, dados devem ser:
    - Estruturados
    - Representativos
    - De boa qualidade
- Possuem um viés
  - Tendência a privilegiar uma ou mais hipóteses que atendam a um dado critério

# Viés indutivo

- Algoritmos de AM precisam ter um viés indutivo
  - Necessário para restringir o espaço de busca
  - Sem viés não há generalização
  - Sem generalização não ocorre aprendizado
    - Modelos (regras / equações) seriam especializados para os dados usados para sua indução

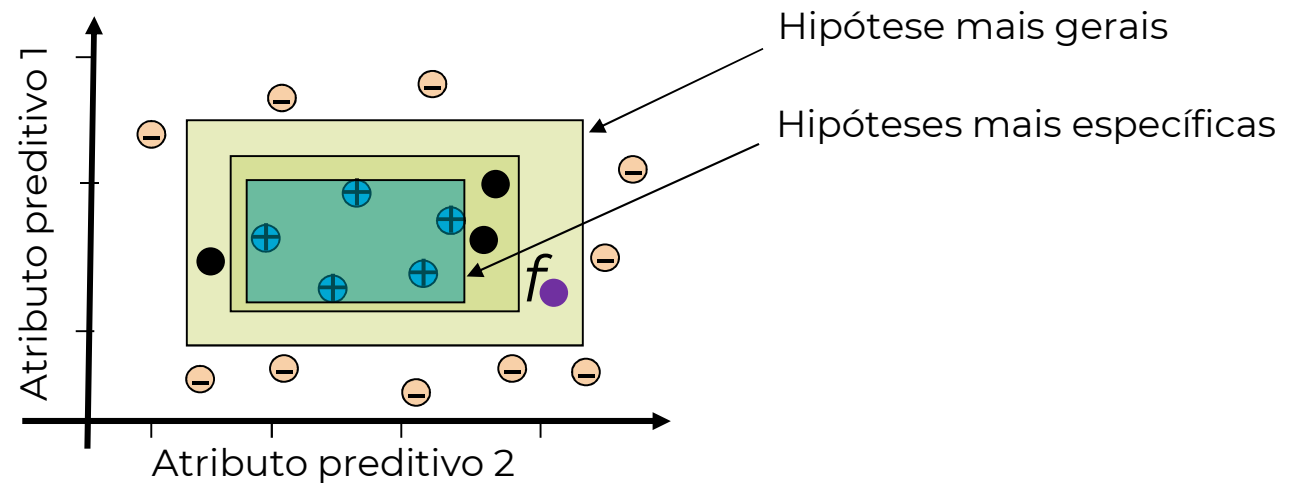


# Viés indutivo

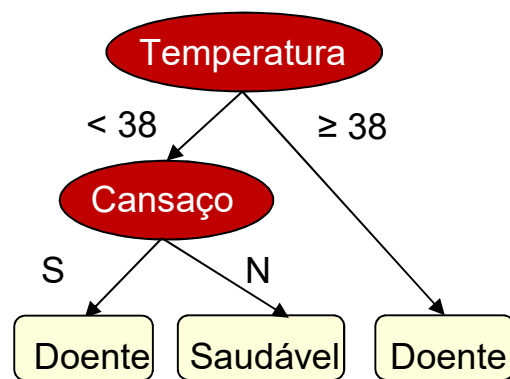
- Viés de preferência ou busca
  - Define como as hipóteses são pesquisadas no espaço de possíveis hipóteses
  - Preferência de algumas hipóteses sobre outras
    - Ex.: preferência por hipóteses simples (curtas)
- Viés de representação ou linguagem
  - Define o espaço de busca de hipóteses
  - Restringe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

# Viés de busca

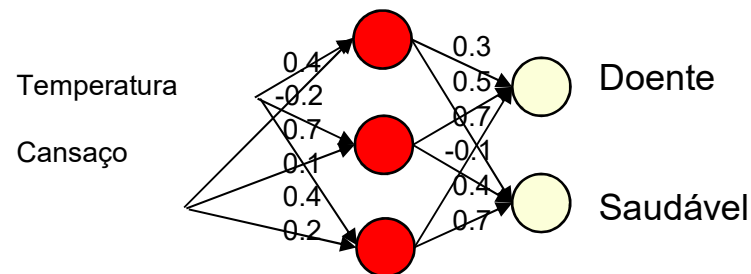
- Seja  $H$  o conjunto de todas as possíveis hipóteses que um algoritmo pode encontrar
  - Supor que cada retângulo define um conjunto de hipótese  $h \in H$  e que  $f$  é a hipótese (função) verdadeira



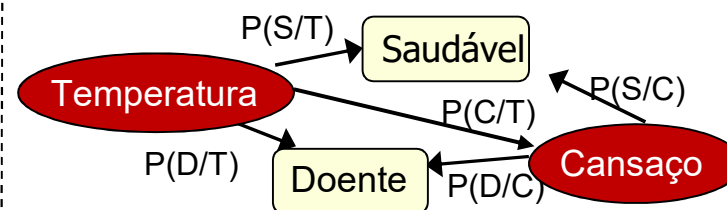
# Viés de representação



Árvore de decisão



Redes neurais

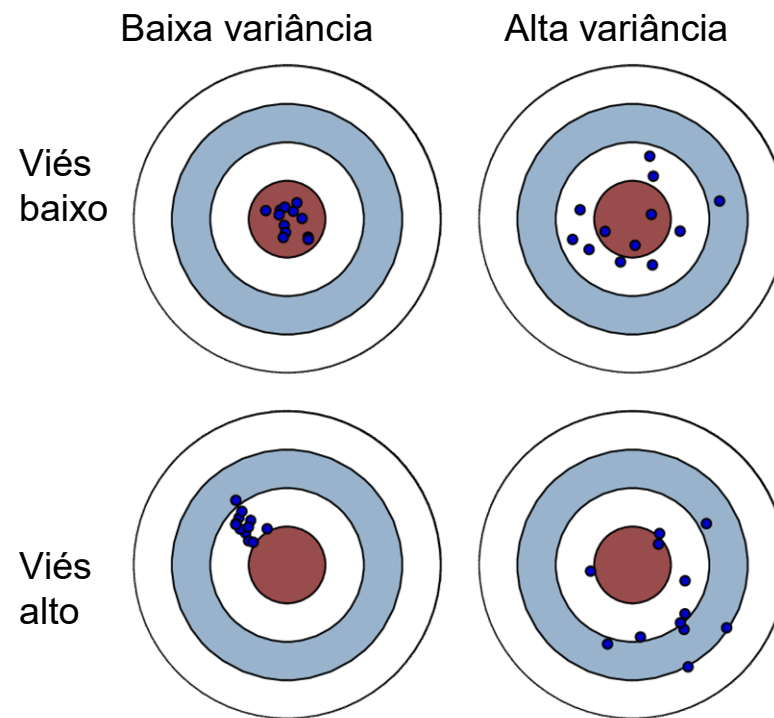


Redes Bayesianas

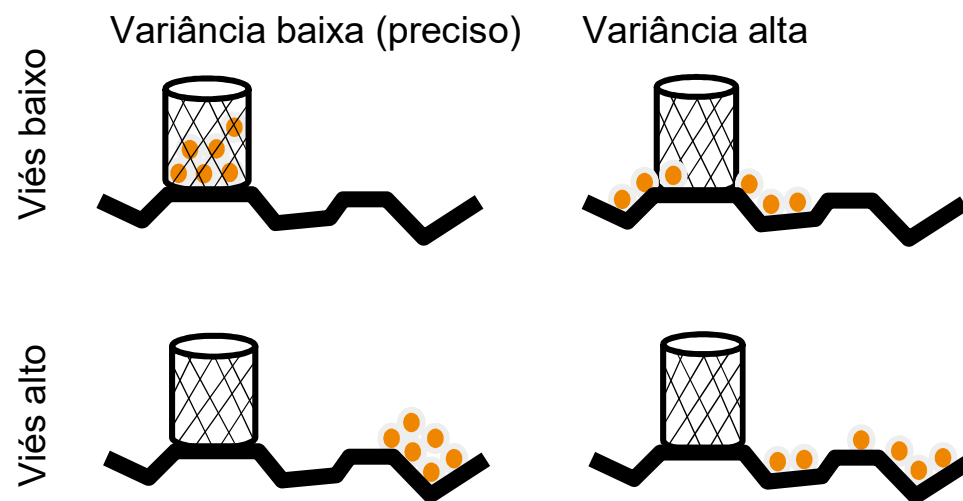
# Viés, Variância e Algoritmos de AM

- Fontes de erro de modelos preditivos
  - Viés (quanto mais alto, pior)
    - Quando algoritmo não aprende um modelo adequado (quanto maior o viés, mais simples a hipótese induzida)
    - Associado a *underfitting*
  - Variância (quanto mais alto, pior)
    - Quando algoritmo presta atenção a detalhes sem importância (alta quando pequenas mudanças nos dados de treinamento alteram o modelo gerado)
    - Associada a *overfitting*
- Precisam ser reduzidos

# Viés e Variância

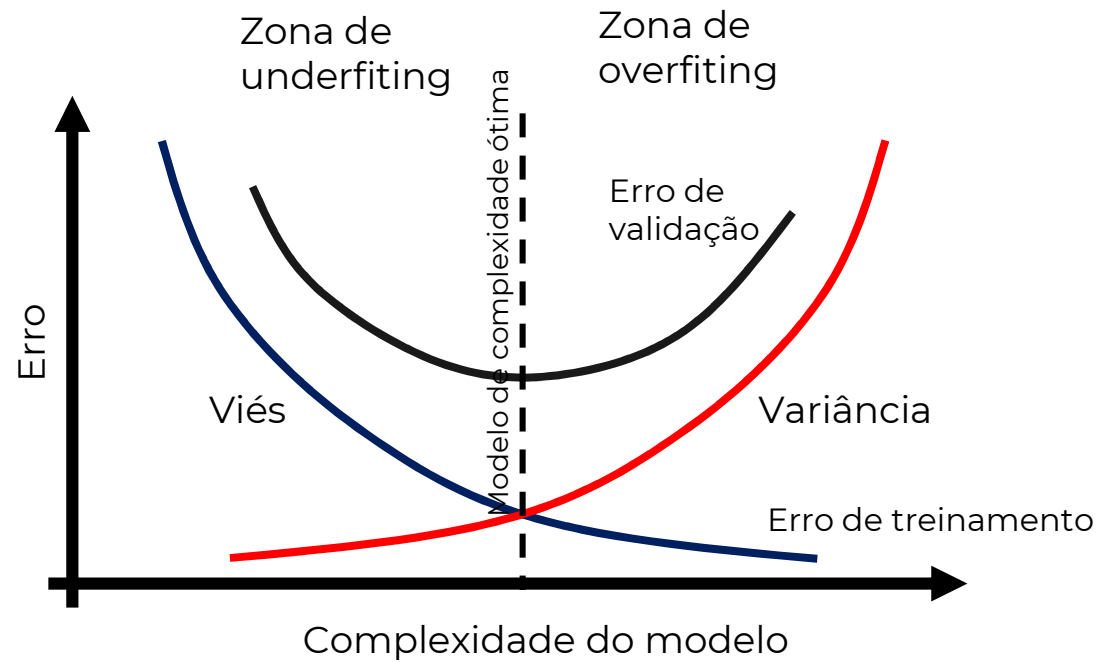


# Viés e Variância

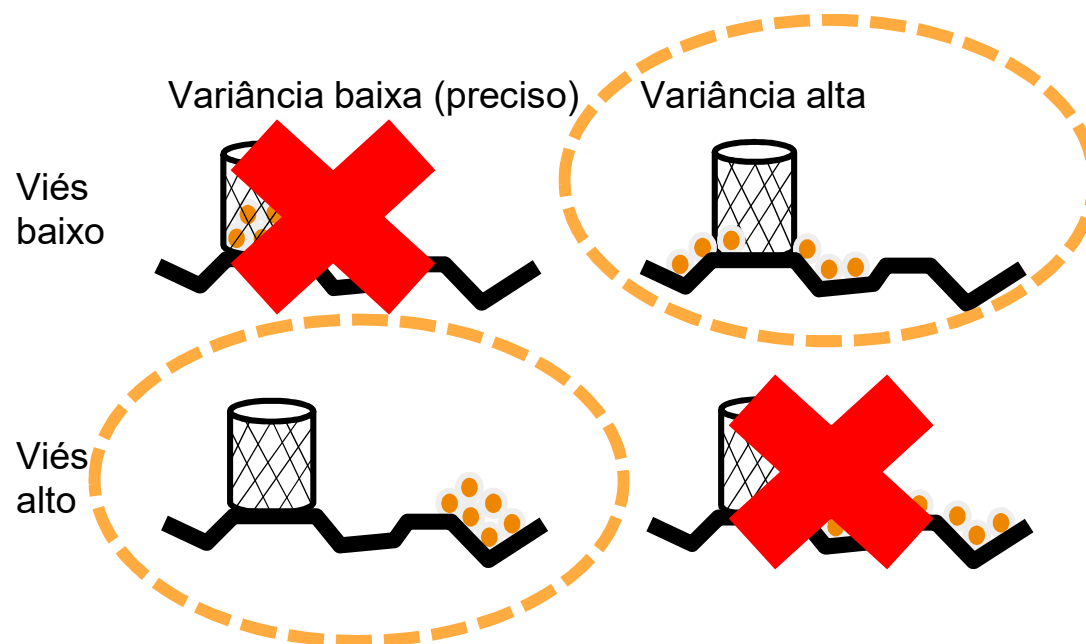




# Viés-Variância trade-off



# Viés e Variância



# Algoritmos preditivos

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

Geométricos

# Ajuste de (hiper-)parâmetros

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

# Ajuste de hiper-parâmetros

- Desempenho dos modelos induzidos por um algoritmo depende dos valores dos hiper-parâmetros do algoritmo
- 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
- Algoritmos de Aprendizado de Máquina
- Viés indutivo
- Dilema viés-variância
- Algoritmos preditivos
- Parâmetros e hiper-parâmetros

Fim do  
apresentação