

Aprendizado de Máquina

Aprendizado



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

Tópicos

- Aprendizado de Máquina
- Viés indutivo
- Algoritmos de Aprendizado de Máquina
- Dilema viés-variância
- Algoritmos de classificação

© André de Carvalho - ICMC/USP 2

Algoritmos de AM

- Induzem modelos (hipóteses) a partir de um conjunto de dados
- Dados precisam
 - Ser estruturados
 - Ter boa qualidade
 - Ser representativos
- Algoritmos de AM indutivo possuem um viés
 - Tendência a privilegiar uma dada hipótese ou conjunto de hipóteses

© André de Carvalho - ICMC/USP 3

Viés indutivo

- Algoritmos de AM precisam ter um viés indutivo
 - Necessário para restringir o espaço de busca
 - Se não houvesse viés, não haveria generalização
 - Regras / equações seriam especializados para os dados usados para a indução modelos
 - Dados de treinamento

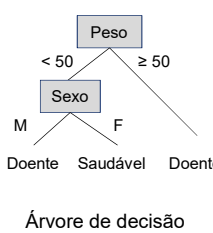
© André de Carvalho - ICMC/USP 4

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 ADs

© André de Carvalho - ICMC/USP 5

Viés de representação



Árvore de decisão

0.45	-0.40	0.54	0.12	0.98	0.37
-0.45	0.11	0.91	0.34	-0.20	0.83
-0.29	0.32	-0.25	-0.51	0.41	0.70

Redes neurais

Se Peso ≥ 50 então Doente
 Se Peso < 50 e Sexo = M então Doente
 Se Peso < 50 e Sexo = F então Saudável

Conjunto de regras

© André de Carvalho - ICMC/USP 6

Viés de representação

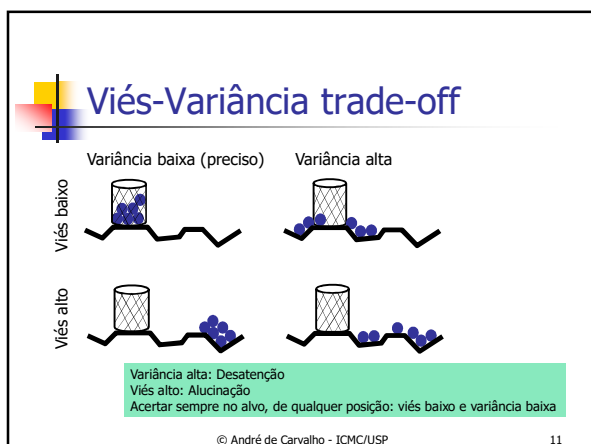
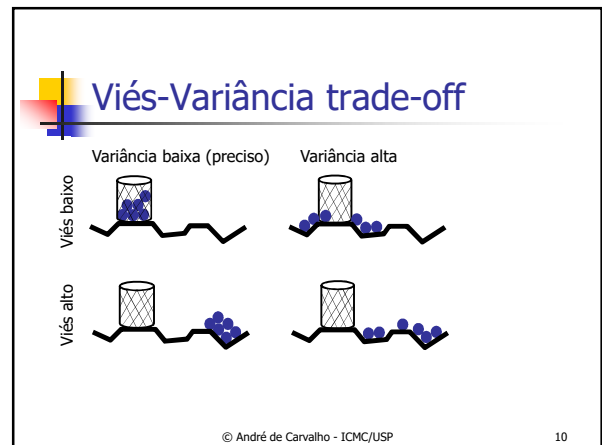
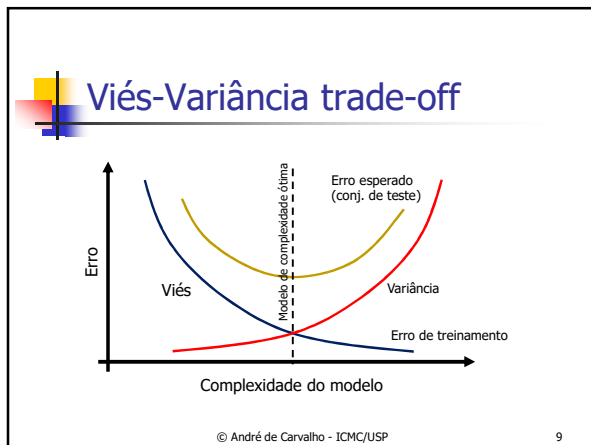
- Influência da representação no viés
 - Ex.: Hipótese é representada por uma regra que apresenta apenas conjunções
 - Regras com negações ou disjunções não fazem parte do espaço de hipóteses
 - Reduz consideravelmente espaço de possíveis hipóteses

© André de Carvalho - ICMC/USP 7

Algoritmos de AM

- Fontes de erro de algoritmos AM
 - Viés
 - Quando algoritmo aprende um modelo incorreto
 - Associado a *underfitting*
 - Variância
 - Quando algoritmo presta atenção a detalhes sem importância
 - Associado a *overfitting*
- Precisam ser reduzidos

© André de Carvalho - ICMC/USP 8



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

© André de Carvalho - ICMC/USP 12

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
 - Alucinação

© André de Carvalho - ICMC/USP

13

Bom algoritmo de AM

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



© André de Carvalho - ICMC/USP

14

Classificação

- Quando função para definir fronteira de decisão se torna mais 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

© André de Carvalho - ICMC/USP

15

Classificação

- Sintomas que poderiam permitir um melhor modelo para diagnóstico:
 - Batimentos cardíacos
 - Idade
 - Peso
 - Pressão
 - Temperatura
 - Taxas em uma amostra de sangue

© André de Carvalho - ICMC/USP

16

Classificação

- 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

© André de Carvalho - ICMC/USP

17

Algoritmos de classificação

- Centenas de novos a cada ano
- Principais
 - Indução de Árvores de Decisão
 - Indução de conjuntos de regras
 - Redes Neurais
 - Máquinas de Vetores de Suporte
 - K-NN
 - Regressão Logística
 - Redes Bayesianas

© André de Carvalho - ICMC/USP

18

Algoritmos de classificação

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

© André de Carvalho - ICMC/USP 19

Algoritmos de classificação

- 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

© André de Carvalho - ICMC/USP 20

Algoritmos de AM

- Desempenho depende de valores de seus hiper-parâmetros
- Quanto menos hiper-parâmetros, melhor
- Com muitos hiper-parâmetros, tudo é possível

Com 4 parâmetros eu posso modelar um elefante, e com 5 eu posso fazer ele mover seu tronco

John Von Neuman


© André de Carvalho - ICMC/USP 21

Conclusão

- Aprendizado de Máquina
 - Viés indutivo
 - Algoritmos de Aprendizado de Máquina
 - Dilema viés-variância
 - Algoritmos de classificação

© André de Carvalho - ICMC/USP 22

Perguntas



© André de Carvalho - ICMC/USP 23