

# MAC425/5739 - Inteligência Artificial

Vítor Kei Taira Tamada - 8516250

## Lista de exercícios 4

### Exercício 1

a) Um motivo para um agente inteligente utilizar algoritmos de aprendizagem é o fato de o programador do agente não saber ou ser capaz de prever todas as situações e mudanças possíveis que o agente enfrentará ao longo do tempo. Outro motivo é o programador não saber como programar a solução para certos casos. Em ambas as situações, o agente não teria acesso a todas as informações de antemão.

Se algum desses problemas surgir, os outros algoritmos (busca, lógica e MDP) se tornam muito difíceis se não impossíveis de se implementar pois exigem conhecimento do ambiente, ações, modelos de probabilidade, solução, dentre outras informações.

b) A principal diferença entre aprendizado supervisionado e aprendizado não-supervisionado é que, apesar de ambos buscarem uma função  $Y = f(X) + \epsilon$  que aproxime  $f(X)$ , o supervisionado recebe um conjunto de pontos  $(x_1, y_1), \dots, (x_n, y_n)$  com  $y_i$  dados, enquanto o não-supervisionado não. Ou seja, o aprendizado supervisionado busca aprender as relações, padrões (possui um "professor" que fornece *feedback*) e regularidades enquanto o não-supervisionado precisa descobri-las (não possui tal "professor").

c) Aprendizado por reforço pode ser considerado não-supervisionado, pois, diferente do aprendizado supervisionado, o por reforço não aprende por um conjunto de pares entrada-resposta (*input-response*). Além disso, ao fazer uso de aprendizado por reforço, o agente precisa explorar e experimentar o ambiente para aprender como se comportar nele (inferir o valor de cada ação em cada estado), enquanto, no supervisionado, algumas dessas informações (valores) são fornecidas.

■

### Exercício 2

c) A taxa de erro no conjunto de treinamento não é considerada uma boa medida de desempenho de um classificador, pois essa taxa pode ser baixa devido ao fato de o classificador já ter feito o conjunto de treinamento em questão, o que significa que já é esperado que ele saiba as soluções ou parte delas.

Por outro lado, a taxa de erro no conjunto de testes é considerada uma boa medida, pois é uma indicação de o quão bem o classificador aprendeu sobre o problema.

d) Dado que um modelo é uma representação do problema em questão, temos que:

Viés determina o quão próximo o modelo é do problema: viés baixo significa que o modelo é próximo do problema e vice-versa.

Variância determina a diferença entre os diferentes modelos gerados para o problema: variância baixa significa que os modelos tem pouca diferença entre si e vice-versa.

e) *Overfitting* é um problema que consiste em um classificador tornar-se mais complexo do que o requisitado pelo problema. Por ele ser mais complexo do que o necessário, o custo computacional aumenta e o poder de generalização diminui.

*Underfitting*, por outro lado, é um problema que consiste em um classificador tão simples que acaba tendo performance baixo até mesmo no conjunto de treinamento. Por conta da simplicidade, apesar de exigir pouco do computador, também tem poder de generalização baixo.

■

### Exercício 3

a) Dado que o *Naive Bayes* considera que os atributos são condicionalmente independentes uns dos outros, temos que:

- Atividade = Festa

$$\begin{aligned}
& P(\text{Atividade} = \text{Festa} \mid \text{prazo próximo, sem festa, cansado}) \\
&= \alpha \times P(\text{prazo próximo} \mid \text{Atividade} = \text{Festa}) \times P(\text{sem festa} \mid \text{Atividade} = \text{Festa}) \times P(\text{cansado} \\
&\mid \text{Atividade} = \text{Festa}) \times P(\text{Atividade} = \text{Festa}) \\
&= \alpha \times 2/5 \times 0/5 \times 3/5 \times 5/10 \\
&= 0
\end{aligned}$$

• Atividade = Estudar

$$\begin{aligned}
& P(\text{Atividade} = \text{Estudar} \mid \text{prazo próximo, sem festa, cansado}) \\
&= \alpha \times P(\text{prazo próximo} \mid \text{Atividade} = \text{Estudar}) \times P(\text{sem festa} \mid \text{Atividade} = \text{Estudar}) \times \\
&P(\text{cansado} \mid \text{Atividade} = \text{Estudar}) \times P(\text{Atividade} = \text{Estudar}) \\
&= \alpha \times 1/3 \times 1 \times 1/3 \times 3/10 \\
&= \alpha \times 1/30
\end{aligned}$$

• Atividade = Bar

$$\begin{aligned}
& P(\text{Atividade} = \text{Bar} \mid \text{prazo próximo, sem festa, cansado}) \\
&= \alpha \times P(\text{prazo próximo} \mid \text{Atividade} = \text{Bar}) \times P(\text{sem festa} \mid \text{Atividade} = \text{Bar}) \times P(\text{cansado} \mid \\
&\text{Atividade} = \text{Bar}) \times P(\text{Atividade} = \text{Bar}) \\
&= \alpha \times 0 \times 1 \times 1 \times 1/10 \\
&= 0
\end{aligned}$$

• Atividade = TV

$$\begin{aligned}
& P(\text{Atividade} = \text{TV} \mid \text{prazo próximo, sem festa, cansado}) \\
&= \alpha \times P(\text{prazo próximo} \mid \text{Atividade} = \text{TV}) \times P(\text{sem festa} \mid \text{Atividade} = \text{TV}) \times P(\text{cansado} \mid \\
&\text{Atividade} = \text{TV}) \times P(\text{Atividade} = \text{TV}) \\
&= \alpha \times 1 \times 1 \times 1 \times 1/10 \\
&= \alpha \times 1/10
\end{aligned}$$

**Nota:**  $\alpha = (P(\text{prazo próximo}) \times P(\text{sem festa}) \times P(\text{cansado}))^{-1}$

Portanto, a entrada "prazo próximo, sem festa, cansado" classifica-se como TV, pois é a que tem a maior probabilidade.

b) A entropia de uma variável aleatória  $V$  de valores  $v_k$  é dada por:

$$H(V) = -\sum_k P(v_k) \log_2 P(v_k)$$

Dados atributos

- Prazo = {Urgente, Próximo, Nenhum}
- Festa = {Sim, Não}
- Cansado = {Sim, Não}

e o objetivo

- Atividade = {Festa, Estudar, Bar, TV}

temos as seguintes probabilidades para cada valor de cada atributo dado o conjunto de dados inicial:

Prazo

$$\begin{aligned}
& P(\text{Prazo} = \text{Urgente}) = 3/10 = 0.3 \\
& P(\text{Prazo} = \text{Próximo}) = 4/10 = 0.4 \\
& P(\text{Prazo} = \text{Nenhum}) = 3/10 = 0.3
\end{aligned}$$

Festa

$$\begin{aligned}
& P(\text{Festa} = \text{Sim}) = 5/10 = 0.5 \\
& P(\text{Festa} = \text{Não}) = 5/10 = 0.5
\end{aligned}$$

Cansado

$$\begin{aligned}
& P(\text{Cansado} = \text{Sim}) = 6/10 = 0.6 \\
& P(\text{Cansado} = \text{Não}) = 4/10 = 0.4
\end{aligned}$$

e as seguintes probabilidades para o objetivo

Atividade

$$\begin{aligned}
& P(\text{Atividade} = \text{Festa}) = 5/10 = 0.5 \\
& P(\text{Atividade} = \text{Estudar}) = 3/10 = 0.3 \\
& P(\text{Atividade} = \text{Bar}) = 1/10 = 0.1
\end{aligned}$$

$$P(\text{Atividade} = \text{TV}) = 1/10 = 0.1$$

Logo, a entropia de cada atributo é:

Prazo

$$H_0(\text{Prazo}) = -(0.3 \times \log_2 0.3 + 0.4 \times \log_2 0.4 + 0.3 \times \log_2 0.3) = 1.571$$

Festa

$$H_0(\text{Festa}) = -(0.5 \times \log_2 0.5 + 0.5 \times \log_2 0.5) = 1$$

Cansado

$$H_0(\text{Cansado}) = -(0.6 \times \log_2 0.6 + 0.4 \times \log_2 0.4) = 0.971$$

e do objetivo é

$$H_0(\text{Atividade}) = -(0.5 \times \log_2 0.5 + 0.3 \times \log_2 0.3 + 0.1 \times \log_2 0.1 + 0.1 \times \log_2 0.1) = 1.686$$

Portanto, o ganho de informação inicial para cada atributo é:

Prazo

$$\text{Gain}_0(\text{Prazo}) = H_0(\text{Atividade}) - H_0(\text{Prazo}) = 1.686 - 1.571 = 0.115$$

Festa

$$\text{Gain}_0(\text{Festa}) = H_0(\text{Atividade}) - H_0(\text{Festa}) = 1.686 - 1 = 0.686$$

Cansado

$$\text{Gain}_0(\text{Cansado}) = H_0(\text{Atividade}) - H_0(\text{Cansado}) = 1.686 - 0.971 = 0.715$$

Logo, o primeiro nó da árvore (raiz) é o atributo "Cansado".

Recalculando as probabilidades, temos:

Cansado = Sim

• Prazo

$$P(\text{Prazo} = \text{Urgente}) = 2/6$$

$$P(\text{Prazo} = \text{Próximo}) = 3/6 = 0.5$$

$$P(\text{Prazo} = \text{Nenhum}) = 1/6$$

• Festa

$$P(\text{Festa} = \text{Sim}) = 3/6 = 0.5$$

$$P(\text{Festa} = \text{Não}) = 3/6 = 0.5$$

• Atividade

$$P(\text{Atividade} = \text{Festa}) = 3/6 = 0.5$$

$$P(\text{Atividade} = \text{Estudar}) = 1/6$$

$$P(\text{Atividade} = \text{Bar}) = 1/6$$

$$P(\text{Atividade} = \text{TV}) = 1/6$$

Cansado = Não

• Prazo

$$P(\text{Prazo} = \text{Urgente}) = 1/4 = 0.25$$

$$P(\text{Prazo} = \text{Próximo}) = 1/4 = 0.25$$

$$P(\text{Prazo} = \text{Nenhum}) = 2/4 = 0.5$$

• Festa

$$P(\text{Festa} = \text{Sim}) = 2/4 = 0.5$$

$$P(\text{Festa} = \text{Não}) = 2/4 = 0.5$$

• Atividade

$$P(\text{Atividade} = \text{Festa}) = 2/4 = 0.5$$

$$P(\text{Atividade} = \text{Estudar}) = 2/4 = 0.5$$

$$P(\text{Atividade} = \text{Bar}) = 0/6 = 0$$

$$P(\text{Atividade} = \text{TV}) = 0/6 = 0$$

Recalculando as entropias, temos:

Cansado = Sim

Prazo

$$H_1(Prazo) = -(2/6 \times \log_2(2/6) + 0.5 \times \log_2 0.5 + 1/6 \times \log_2(1/6)) = 1.459$$

Festa

$$H_1(Festa) = -(0.5 \times \log_2 0.5 + 0.5 \times \log_2 0.5) = 1$$

Atividade

$$H_1(Atividade) = -(0.5 \times \log_2 0.5 + 1/6 \times \log_2(1/6) + 1/6 \times \log_2(1/6) + 1/6 \times \log_2(1/6)) = 1.793$$

Portanto, o ganho de informação desta etapa para cada atributo é:

Prazo

$$Gain_1(Prazo) = H_1(Atividade) - H_1(Prazo) = 1.793 - 1.459 = 0.334$$

Festa

$$Gain_1(Festa) = H_1(Atividade) - H_1(Festa) = 1.793 - 1 = 0.793$$

Logo, o segundo nó da árvore (raíz) é o atributo "Festa".

Cansado = Não

Prazo

$$H_1(Prazo) = -(1/4 \times \log_2(1/4) + 1/4 \times \log_2(1/4) + 2/4 \times \log_2(2/4)) = 1.5$$

Festa

$$H_1(Festa) = -(0.5 \times \log_2 0.5 + 0.5 \times \log_2 0.5) = 1$$

Atividade

$$H_1(Atividade) = -(0.5 \times \log_2 0.5 + 0.5 \times \log_2(0.5) + 0 \times \log_2 0 + 0 \times \log_2 0) = 1$$

Portanto, o ganho de informação desta etapa para cada atributo é:

Prazo

$$Gain_1(Prazo) = H_1(Atividade) - H_1(Prazo) = 1 - 1.5 = -0.5$$

Festa

$$Gain_1(Festa) = H_1(Atividade) - H_1(Festa) = 1 - 1 = 0$$

Logo, o segundo nó da árvore (raíz) é o atributo "Festa".

Recalculando as probabilidades, temos:

Cansado = Sim, Festa = Sim

• Prazo

$$P(\text{Prazo} = \text{Urgente}) = 1/3$$

$$P(\text{Prazo} = \text{Próximo}) = 2/3$$

$$P(\text{Prazo} = \text{Nenhum}) = 0/3 = 0$$

• Atividade

$$P(\text{Atividade} = \text{Festa}) = 3/3 = 1$$

$$P(\text{Atividade} = \text{Estudar}) = 0/3 = 0$$

$$P(\text{Atividade} = \text{Bar}) = 0/3 = 0$$

$$P(\text{Atividade} = \text{TV}) = 0/3 = 0$$

Cansado = Sim, Festa = Não

• Prazo

$$P(\text{Prazo} = \text{Urgente}) = 1/3$$

$$P(\text{Prazo} = \text{Próximo}) = 1/3$$

$$P(\text{Prazo} = \text{Nenhum}) = 1/3$$

• Atividade

$$P(\text{Atividade} = \text{Festa}) = 0/3 = 0$$

$$P(\text{Atividade} = \text{Estudar}) = 1/3$$

$$P(\text{Atividade} = \text{Bar}) = 1/3$$

$$P(\text{Atividade} = \text{TV}) = 1/3$$

Cansado = Não, Festa = Sim

- Prazo

$$P(\text{Prazo} = \text{Urgente}) = 0/2$$

$$P(\text{Prazo} = \text{Próximo}) = 0/2$$

$$P(\text{Prazo} = \text{Nenhum}) = 2/2 = 1$$

- Atividade

$$P(\text{Atividade} = \text{Festa}) = 2/2 = 1$$

$$P(\text{Atividade} = \text{Estudar}) = 0/2 = 0$$

$$P(\text{Atividade} = \text{Bar}) = 0/2 = 0$$

$$P(\text{Atividade} = \text{TV}) = 0/2 = 0$$

Cansado = Não, Festa = Não

- Prazo

$$P(\text{Prazo} = \text{Urgente}) = 1/2 = 0.5$$

$$P(\text{Prazo} = \text{Próximo}) = 1/2 = 0.5$$

$$P(\text{Prazo} = \text{Nenhum}) = 0/2 = 0$$

- Atividade

$$P(\text{Atividade} = \text{Festa}) = 0/2 = 0$$

$$P(\text{Atividade} = \text{Estudar}) = 2/2 = 1$$

$$P(\text{Atividade} = \text{Bar}) = 0/2 = 0$$

$$P(\text{Atividade} = \text{TV}) = 0/2 = 0$$

Como há apenas mais um atributo, não é necessário calcular sua entropia e ganho de informação, pois ele determinará qual a atividade.

**Esquema da árvore de decisão:**