

Lista de Exercícios

Tratamento de Incerteza baseado em Probabilidade

- 1) Explique o termo “probabilidade subjetiva” no contexto de um agente que raciocina sobre incerteza baseando em probabilidade.
- 2) Explique o que é uma variável randômica e como elas podem ser classificadas de acordo com o domínio.
- 3) Seja f uma proposição, e x um valor do domínio de f . Seja y um valor real entre 0 e 1. Explique o que significa $P(f=x) = y$, de acordo com a semântica dos mundos possíveis.
- 4) Seja h e e proposições. O que é $P(h | e)$. O que esta probabilidade significa de acordo com a semântica dos mundos possíveis?
- 5) Explique por que o Teorema de Bayes é interessante. Por que a probabilidade condicional pode estar disponível em um sentido, mas não em outro?
- 6) Considere a situação na qual queremos raciocinar sobre o tempo e as viagens. Usaremos três variáveis randômicas booleanas para representar “o tempo está ruim” (T), “os vôos entre Maringá e Curitiba estão canceladas” (V) e “as viagens de carro na BR369, entre Maringá e Curitiba, estão sendo atrasadas” (B).

Intuitivamente, sabemos que, para uma dada condição de tempo, os vôos cancelados não afetam atrasos nas BRs, e que atrasos nas BRs não afetam os cancelamento de vôos. Da experiência, sabemos que, se o tempo está ruim, existem 80% de chance que os vôos sejam cancelados. Se o tempo está bom, existem 40% de chances que os vôos sejam cancelados. Além do mais, o tempo ruim causa atrasos nas BRs 30% das vezes, mas com tempo bom as chances de atrasos são de somente 10%. Existem 20% de chance de tempo ruim em Maringá nesta época do ano.

- a. Crie uma rede Bayseana para este problema, inclua seu grafo e suas tabelas de distribuição condicional.
 - b. Calcule $P(\neg V, T, B)$.
- 7) Imagine que 99% do tempo a doença *RE* (DRE) cause olhos vermelhos nas pessoas que a contraem. No entanto, somente 2% de todas as pessoas têm olhos vermelhos no dia-a-dia, e somente 1% contraem a doença DRE. Você tem olhos vermelhos. Qual é a probabilidade de você ter DRE?

Lógica Fuzzy

- 8) Explique a diferença entre o tratamento de incerteza baseado em probabilidade e o tratamento de incerteza baseado em lógica *fuzzy*. Explique os termos “graus de crença” e “graus de verdade”.
- 9) Enumere as diferenças entre a teoria clássica dos conjuntos (*crisp*) e a teoria dos conjuntos *fuzzy*.
- 10) O que é uma variável lingüística no contexto da lógica *fuzzy* e como ela pode ser representada?
- 11) Quais são os principais componentes que formam um sistema especialista *fuzzy*? Explique rapidamente do que se trata cada um deles.
- 12) Explique quais são as fases de um sistema *fuzzy* e do que se trata cada delas.

- 13) O que significa dizer que uma regra é disparada na fase de processamento? Quanto o conseqüente desta regra influenciará na saída do sistema *fuzzy*?
- 14) Explique, sucintamente, o processo de construção de um sistema especialista baseado em lógica *fuzzy*.
- 15) Supõe que uma universidade moderna resolveu mudar o seu método de avaliação, por pensar que o método em que um aluno com média 6,0 é aprovado e um aluno com média 5,9 é reprovado é um método injusto. A universidade resolveu então, que, além da média anual na disciplina, outros méritos devem ser avaliados para a aprovação ou não de um aluno, como por exemplo, frequência, e participação em sala de aula. Desta forma, pede-se que você formule um sistema *fuzzy* que leve em consideração estes méritos, e dê como resposta se o aluno deve ser aprovado, reprovado ou fazer um curso de férias para tentar melhorar seu aproveitamento. Desta forma, defina os itens abaixo, sempre fundamentando sua resposta:
- Os conjuntos *fuzzy* (no mínimo três) e suas relações de pertinência para as variáveis de entrada média_anual, frequência e participação. Justifique a sua escolha para a forma da função de pertinência (trapezoidal, triangular, gaussiana...);
 - Os conjuntos *fuzzy* (no mínimo três) e suas relações de pertinência para a variável de saída situação_aluno. Justifique a sua escolha para a forma da função de pertinência (trapezoidal, triangular, gaussiana...);
 - As regras de funcionamento do sistema;
 - O critério para definir o nível de pertinência do conseqüente quando o antecedente de uma regra com mais de uma variável de entrada disparar;
 - O método de *defuzzyficação* utilizado.

Aprendizagem

- 16) Quais questões afetam o projeto do elemento de aprendizagem de um agente? Explique cada um deles.
- 17) Os tipos de aprendizagem são classificados de acordo com o tipo de realimentação que o elemento de aprendizagem terá à sua disposição. Quais são estes tipos e dê exemplos de cada um deles.
- 18) A aprendizagem não supervisionada pode determinar sozinha se uma situação é boa ou ruim para um agente? Por quê?
- 19) Explique qual é o objetivo da aprendizagem indutiva. Do que precisamos para “aprender” uma hipótese h ?
- 20) Uma hipótese h induzida por um método de aprendizagem indutiva pode ser facilmente validada? Por quê?
- 21) O que é o princípio da “Navalha de Ockham”? e como ele é aplicado no algoritmo de aprendizagem por árvore de decisão?
- 22) Explique rapidamente como funciona o algoritmo de aprendizagem para árvores de decisão visto em sala de aula.
- 23) O que significa dizer que um atributo é inútil? E que um atributo é perfeito?
- 24) De acordo com o conjunto de treinamento abaixo, aplique o algoritmo de aprendizagem em árvores de decisão para derivar uma árvore decisão condizente com os exemplos.

Atributos					Predicado meta
Dia	Perspectiva	Temperatura	Umidade	Vento	JogarTennis
D1	Ensolarado	Quente	Alta	Fraco	Não
D2	Ensolarado	Quente	Alta	Forte	Não
D3	Nublado	Quente	Alta	Fraco	Sim
D4	Chuvoso	Moderada	Alta	Fraco	Sim
D5	Chuvoso	Fresca	Normal	Fraco	Sim
D6	Chuvoso	Fresca	Normal	Forte	Não
D7	Nublado	Fresca	Normal	Forte	Sim
D8	Ensolarado	Moderada	Alta	Fraco	Não
D9	Ensolarado	Fresca	Normal	Fraco	Sim
D10	Chuvoso	Moderada	Normal	Fraco	Sim
D11	Ensolarado	Moderada	Normal	Forte	Sim
D12	Nublado	Moderada	Alta	Forte	Sim
D13	Nublado	Quente	Normal	Fraco	Sim
D14	Chuvoso	Moderada	Alta	Forte	Não

25) Assuma que lhe é dado o conjunto rotulado *exemplos de treinamento* abaixo, onde cada característica tem três (3) possíveis valores *a*, *b*, ou *c*. Aplique o algoritmo de árvore de decisão a este conjunto.

<i>F1</i>	<i>F2</i>	<i>F3</i>	<i>Classificação</i>
<i>a</i>	<i>a</i>	<i>a</i>	+
<i>b</i>	<i>b</i>	<i>c</i>	-
<i>a</i>	<i>b</i>	<i>c</i>	-
<i>c</i>	<i>b</i>	<i>c</i>	+
<i>c</i>	<i>a</i>	<i>c</i>	+
<i>b</i>	<i>a</i>	<i>a</i>	-

26) Assuma que você receba os seguintes atributos com os possíveis valores apresentados: $F1 \in \{A, B, C\}$; $F2 \in \{D, E\}$; $F3 \in \{G, H, I\}$; e $F4 \in \{J, K\}$. Considere os seguintes exemplos de treinamento.

ex1	$F1 = A$	$F2 = E$	$F3 = H$	$F4 = J$	categoria = +
ex2	$F1 = A$	$F2 = E$	$F3 = G$	$F4 = K$	categoria = -
ex3	$F1 = A$	$F2 = D$	$F3 = G$	$F4 = J$	categoria = +

ex4	F1 = B	F2 = D	F3 = G F4 = J	categoria = +
ex5	F1 = B	F2 = E	F3 = H F4 = J	categoria = +
ex6	F1 = B	F2 = E	F3 = H F4 = K	categoria = -
ex7	F1 = C	F2 = D	F3 = H F4 = K	categoria = -
ex8	F1 = C	F2 = D	F3 = G F4 = J	categoria = -

A. Qual escore a formula de *ganho de informação* atribuí para cada um dos três atributos? (Estes cálculos podem ser úteis: $I(x,y)=I(y,x)$, $I(1,0)=0$, $I(1/2,1/2)=1$, $I(1/3,2/3)=0.92$, $I(1/4,3/4)=0.81$, $I(1/5,4/5)=0.72$, $I(1/6,5/6)=0.65$, $I(1/7,6/7)=0.59$, $I(1/8,7/8)=0.54$). **Tome cuidado de mostrar todo seu desenvolvimento.**

B. Que atributo seria escolhido como raiz? (Resolva os empates em favor de *F1* sobre *F2* sobre *F3* sobre *F4* e '+' sobre '-').

C. Mostre o próximo nó *interior*, se existir algum, que o algoritmo de aprendizado de árvore de decisão selecionaria. Note que, para evitar gastar tempo excessivo, está sendo pedido que você crie no máximo os dois (2) primeiros nós interiores. Crie o maior número de nós folhas possíveis, dada a sua escolha de nós interiores. Novamente, **tome o cuidado de mostrar todo seu desenvolvimento.**

27) Considere coletar dados aonde cada exemplo tem três atributos X_1 , X_2 e X_3 , e onde X_1 possui valores A e B, X_2 possui valores C e D, e X_3 possui valores E, F, G. Seu objetivo é construir uma árvore de decisão para um conceito booleano dado pelos seguintes exemplos:

X_1	X_2	X_3	classificação
A	C	G	+
A	D	F	+
A	C	F	-
A	D	F	+
B	D	G	-
B	D	G	-
B	C	G	+
B	C	F	-

(a) Calcule o ganho de informação para cada um dos três atributos X_1 , X_2 e X_3 . Mostre seus cálculos. Saiba que: $I(1,0)=0$, $I(0.5,0.5)=1$, $I(0.75, 0.25)=0.81$, e $I(0.875, 0.125)=0.54$.

(b) Qual atributo o algoritmo de aprendizado da árvore de decisão escolherá primeiro para raiz quando se usa o critério de ganho-máximo? E qual serão as próximas raízes.

28) Considere os seguintes exemplos de treinamento onde quatro exemplo de valores booleanos (X_1 a X_4) são usados para prever um valor booleano (Y_1):

X_1	X_2	X_3	X_4	Y_1
0	1	1	0	1
1	0	1	0	0
1	0	0	1	1
0	1	1	1	0

Desenhe um *perceptron* para esta tarefa, inicializando todos os pesos e *bias* em 0,1. Mostre como os pesos e os *bias* neste *perceptron* irão ser modificados depois de processar os dois primeiros exemplos acima. Use uma taxa de aprendizagem de 0,5. Mostre todos seus cálculos.

29) Considere um *perceptron* que tem três unidades de entradas com valores reais e duas unidades de saída com uma função de ativação *sigmoidal*. Todos os pesos iniciais e os desvios (*bias*) são iguais a 0,5. Assuma que seu professor tenha dito que a saída deve ser 1 e 0 para os valores de entrada $in_1 = 0,7$; $in_2 = -0,5$ e $in_3 = 0,3$.

A) Mostre como a regra de aprendizado do *perceptron* irá alterar a RNA após processar este exemplo de treinamento. Faça α (a taxa de aprendizagem) ser 0.1, e tenha certeza de ajustar os bias das unidades durante o treinamento.

30) Considere um perceptron que tem duas unidades de entradas com valores reais e uma unidade de saída com uma função de ativação *sinhal*. Suponha os pesos iniciais [0.75; 0.5] e o *bias* de -0.6. Assuma o conjunto de exemplos abaixo. Divida o conjunto em dois, um de treinamento de três exemplos e outro de teste de cinco exemplos e aplique o algoritmo de aprendizagem do *perceptron* até a terceira iteração. Após isto verifique a qualidade do aprendizado obtido. Considere α (a taxa de aprendizagem) igual 0.1, e tenha certeza de ajustar os *bias* das unidades durante o treinamento.

Ex	X1	X2	Saída
1	1.0	1.0	1
2	9.4	6.4	-1
3	2.5	2.1	1
4	8.0	7.7	-1
5	7.9	8.4	-1
6	7.0	7.0	-1
7	2.8	0.8	1
8	1.2	3.0	1

31) Assuma que nós temos uma variável aleatória de diagnóstico (vamos chamá-la de **D**) e duas variáveis de medida (vamos chamá-las de **M1** e **M2**). Para simplificar, assuma que as variáveis têm três valores possíveis (e.g., *low*, *medium*, and *high*) e que **D** seja booleana.

Nós coletamos dados de 300 episódios e achamos o seguinte:

D é *true* 100 vezes e para estes casos:

M1=low 50 vezes, **M1=med** 30 vezes, and **M1 = high** 20 vezes

M2=low 10 vezes, **M2=med** 80 vezes, and **M2 = high** 10 vezes

D é *false* 200 vezes e para estes casos:

M1=low 20 vezes, **M1=med** 80 vezes, and **M1 = high** 100 vezes

M2=low 180 vezes, **M2=med** 10 vezes, and **M2 = high** 10 vezes

Assuma que **M1** e **M2** são *condicionalmente independentes* dado **D**.

i. Mostre como a regra de Bayes pode ser usada para calcular $P(\mathbf{D} \mid \mathbf{M1}, \mathbf{M2})$ conhecido os dados acima e sobre a hipótese acima definida. [faça isto **algebricamente** — i.e., usando equações.]

ii. Em um dos episódios nos achamos **M1=low** e **M2=low**. Qual é o diagnóstico mais provável? De uma justificativa numérica.

32) Você recebe alguns documentos, cada um especificado por um vetor de características (X1, X2, X3), no qual cada componente tem um valor binário. Existem duas possíveis classificações para um documento C= 0 ou 1. É 3 vezes mais provável que um documento esteja na classe 1 do que na classe 0. Você também sabe: $P(X1 = 1 \mid C = 0) = 0.25$, $P(X2 = 1 \mid C = 0) = 0.5$, $P(X3 = 1 \mid C = 0) = 0.4$, $P(X1 = 1 \mid C = 1) = 0.5$, $P(X2 = 1 \mid C = 1) = 0.25$, and $P(X3 = 1 \mid C = 1) = 0.3$.

a) Desenhe uma rede Bayseana (com as TPCs) que representam este problema como um modelo de Bayes ingênuo.

b) Dado um novo documento com um vetor característico (X1 = 0, X2 = 1, X3 = 1), determine a classificação deste documento usando o modelo de Bayes ingênuo representado em (a). Mostre seu trabalho.

c) Quais das seguintes sentenças expressam a pressuposição de independência condicional usada na definição do modelo de Bayes ingênuo?

(i) $P(C | X_1, X_2, X_3) = [P(X_1, X_2, X_3 | C) P(C)] / P(X_1, X_2, X_3)$

(ii) $P(X_1, X_2, X_3 | C) = P(X_1 | C) P(X_2 | C) P(X_3 | C)$

(iii) $P(X_1, X_2, X_3) = P(X_1) P(X_2) P(X_3)$