Inteligência Artificial Lista de exercícios: Aprendizado de Máquina
Prof. Hendrik Macedo

1. Analise o conjunto de dados mostrados na tabela abaixo.

Record	A	B	C	Class
1	0	0	0	+
2	0	0	1	−
3	0	1	1	−
4	0	1	1	−
5	0	0	1	+
6	1	0	1	+
7	1	0	1	−
8	1	0	1	−
9	1	1	1	+
10	1	0	1	+

a) Avalie as probabilidades condicionais para *P(A|+), P(B|+), P(C|+), P(A|-), P(B|-) e P(C|-)*.
b) Use a avaliação de probabilidades condicionais calculadas no item anterior para prever o rótulo de classe de uma amostra de teste (*A=0, B=1, C=0*) usando a abordagem *Naïve Bayes*.

2. O modelo Naïve Bayes tem sido usado com sucesso para classificação automática de spam. Observe a modelagem com "*bag-of-words*" a seguir.
   - Cada email possui um rótulo binário Y com valores em {*spam, ham*}
   - Cada palavra w de um email, não importa onde ela apareça, possui P(W = w | Y), onde W representa um dicionário pré-detrminado. Pontuação é ignorada.
   - Considere um email com K palavras $w_1, \dots, w_k$. Por exemplo, o email "*hi hi you*" possui $w_1 = hi, w_2 = hi, w_3 = you$. Seus rótulos são dados pelo $arg\ max_y\ P(Y = y|w_1, \dots, w_k) = argmax_y P(Y = y) \prod_{i=1}^{K} P(W = w_i\ |Y = y)$.

   a) Você possui um classificador de spam treinado em um grande corpus de e-mails. Abaixo segue uma tabela com algumas probabilidades estimadas.

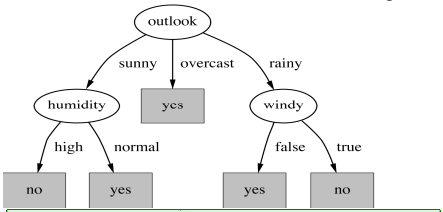
W	note	to	self	become	perfect
$P(W \mid Y = \text{spam})$	1/6	1/8	1/4	1/4	1/8
$P(W \mid Y = \text{ham})$	1/8	1/3	1/4	1/12	1/12

   Você recebe um novo email para classificar, contendo apenas duas palavras: `perfect none.`
   Circule todos os valores de $P(Y = spam)$ para os quais o modelo iria classificar esse novo email como sendo "spam".
   0 0.2 0.4 0.6 0.8 1

   b) Você possui apenas três e-mails como conjunto de treinamento:
   (Spam) `dear sir, I write to you in hope of recovering my gold watch.`
   (Ham) `hey, lunch at 12?`
   (Ham) `fine, watch it tomorrow night.`
   Pinte os círculos correspondentes a valores que você estimaria para as probabilidades dadas abaixo.

$P(W = sir \mid Y = spam)$	0	1/10	1/5	1/3	2/3	Nenhum dos anteriores
$P(W = watch \mid Y = ham)$	0	1/10	1/5	1/3	2/3	Nenhum dos anteriores
$P(W = gauntlet \mid Y = ham)$	0	1/10	1/5	1/3	2/3	Nenhum dos anteriores
$P(Y = ham)$	0	1/10	1/5	1/3	2/3	Nenhum dos anteriores

3. Qual é a taxa de acerto da árvore abaixo para o conjunto de teste que vem a seguir?



Outlook	Temp	Humidity	Windy	Play
Sunny	Cool	High	True	No
Overcast	Hot	High	True	No
Overcast	Hot	High	False	Yes
Rainy	Mild	Normal	False	Yes
Rainy	Cool	Normal	True	Yes

4. Desenhe árvores de decisão que representem os seguintes conceitos (sendo A, B, C e D variáveis booleanas):
   a. A∧B
   b. A∨(B∧C)
   c. A $\otimes$ B
   d. (A∧B)∨(C∧D)

5. Para a tabela ao lado, considere que C é o atributo de classificação, P, Q, e C assumem valores “Y”es ou “N”o e R possui três valores possíveis: 1, 2, e 3. Utilize o algoritmo de aprendizagem ID3 e mostre qual atributo deve ocupar a raiz da árvore de decisão resultante (obs: utilize a ENTROPIA como grandeza de medida de impureza).

P	Q	R	C	Número de instâncias (exemplos) coletadas.
Y	Y	1	N	20
Y	Y	2	Y	1
Y	Y	3	Y	2
Y	N	1	Y	8
Y	N	2	Y	2
Y	N	3	Y	0
N	Y	1	N	12
N	Y	2	Y	2
N	Y	3	Y	1
N	N	1	Y	25
N	N	2	Y	1
N	N	3	Y	4