

1º Teste de Aprendizagem Automática

3 páginas, 5 perguntas, 2 folhas de resposta. Duração: 2 horas
DI, FCT/UNL, 30 de Abril de 2014

Pergunta 1 [4 valores] Foram criadas, por amostragem aleatória com reposição, 200 réplicas de um conjunto de 50 exemplos de valor conhecido. Para cada um de seis modelos de regressão diferentes (A a F) foram calculadas 200 hipóteses ajustando cada modelo a cada uma das réplicas do conjunto inicial. Finalmente, foram calculados os seguintes valores X e Y para cada modelo:

Modelo	X	Y
A	52	0.012
B	13	0.23
C	2	0.51
D	0.4	7
E	0.2	86
F	0.1	2067

$$X = \frac{1}{50} \sum_{t=1}^{50} [\bar{g}(x_t) - y_t]^2 \quad Y = \frac{1}{50 \times 200} \sum_{t=1}^{50} \sum_{i=1}^{200} [\bar{g}(x_t) - g_i(x_t)]^2$$

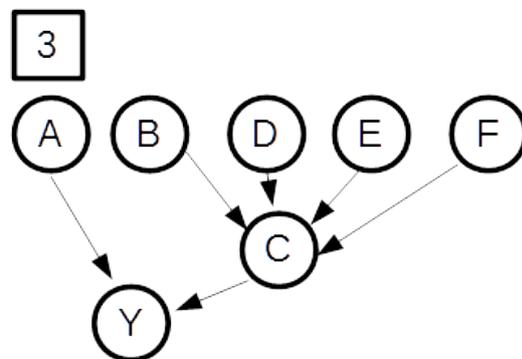
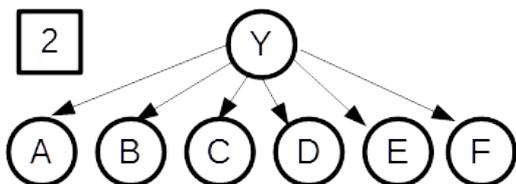
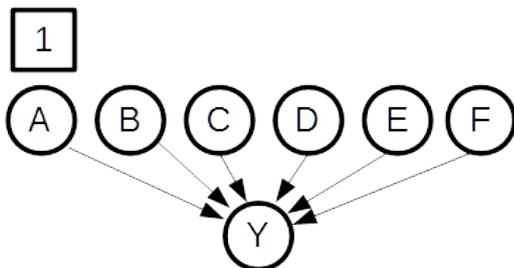
onde t é o índice do exemplo no conjunto inicial, $\bar{g}(x_t)$ é a média dos valores previstos para x_t pelas 200 hipóteses obtidas, y_t o valor real do ponto x_t e $g_i(x_t)$ o valor previsto para o ponto x_t pela hipótese obtida treinando o modelo com a réplica i .

1.a) Indique um modelo cujo erro de generalização se deverá principalmente à incapacidade de se ajustar adequadamente aos exemplos deste conjunto. Justifique a sua resposta.

1.b) Indique um modelo cujo erro de generalização se deverá principalmente ao ajustamento excessivo aos exemplos de treino. Justifique a sua resposta.

1.c) Indique o modelo que terá menor erro de generalização. Justifique a sua resposta.

Pergunta 2 [4 valores] O proprietário de um quiosque, aficionado de redes de Bayes, quer criar um classificador para decidir se concede ou não crédito a clientes que pedem fiado. Ele consegue descrever cada cliente com seis atributos categóricos (A a F) e cada um dos atributos tem quatro categorias. Por exemplo, o atributo A é o veículo que o cliente conduz e tem as categorias de luxo, pesado, utilitário e mota. Ele está indeciso acerca de qual das três redes de Bayes abaixo (1, 2 ou 3) deve usar, onde Y é a classe de cada cliente (fiável ou não fiável).



2.a) Escolha uma das redes (indicando qual escolheu) e escreva a expressão da distribuição de probabilidade conjunta das variáveis nessa rede.

2.b) Considerando que o dono do quiosque já tem um registo de 22 clientes, com todos os atributos, e sabendo que 10 pagaram as dívidas a tempo e 12 não, recomende, justificando, qual das três redes de Bayes será a mais apropriada para prever a fiabilidade dos próximos clientes.

Pergunta 3 [6 valores] Considere o seguinte modelo de classificação onde $g(x)$ é o valor de saída para o exemplo x e os valores w_n são os $M+1$ coeficientes do modelo (contando com w_0 também), sendo M a dimensão dos vectores de entrada:

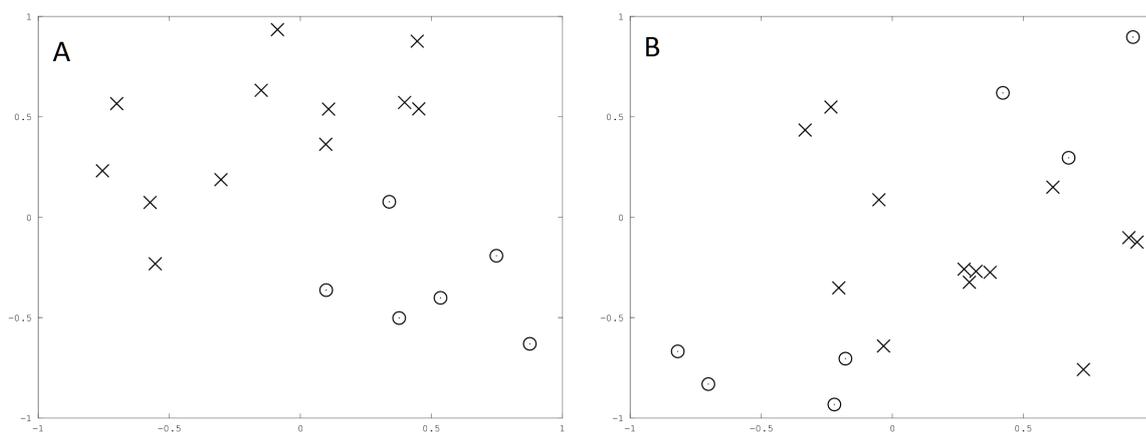
$$g(x) = \frac{1}{1 + e^{-net(x)}} \quad net(x) = w_0 + \sum_{i=1}^M w_i x_i$$

Para obter cada hipótese o modelo é treinado apresentando os exemplos repetidas vezes e por ordem aleatória. De cada vez que se apresenta um exemplo x_t , o modelo é actualizado alterando cada coeficiente w_n da seguinte forma:

$$\Delta w_n = \eta \left(y(x_t) - g(x_t) \right) g(x_t) \left(1 - g(x_t) \right) x_t^n$$

onde $y(x_t)$ é a classe verdadeira do exemplo x_t , que pode ser 0 ou 1, e x_t^n é o valor do atributo de índice n de x_t , considerando este valor igual a 1 se n for 0. O valor η é uma constante que controla o ritmo da aprendizagem. Depois do treino, considera-se que o exemplo x está na classe 1 se $g(x)$ for maior que 0.5, ou na classe 0 caso contrário ($g(x)$ é um valor entre 0 e 1).

Considere estes dois conjuntos de exemplos, A e B, onde um círculo representa a classe 0 e uma cruz a classe 1:



3.a) Indique, justificando, qual dos dois conjuntos de exemplos se pode classificar sem erros com o modelo descrito.

3.b) Explique resumidamente como pode modificar o modelo, usando funções do mesmo tipo, para classificar o outro conjunto que este modelo não consegue classificar sem erros (não precisa de escrever as equações, apenas demonstrar compreender os conceitos).

3.c) Explique também que modificações teria de fazer à função que actualiza os coeficientes se fizesse as alterações pedidas em 3.b (não precisa de escrever as equações, apenas demonstrar compreender os conceitos).

Pergunta 4 [4 valores] Para classificar o conjunto de dados representado na sua folha de resposta, foi treinado um classificador obtendo os valores de α que minimizam uma das seguintes expressões (A ou B) onde N é o número de exemplos, y o valor da classe de cada exemplo (representados com um círculo para a classe -1 e uma cruz para a classe +1) e x o vector com as coordenadas de cada ponto.

$$A) \min_{\alpha} \left(\frac{1}{2} \sum_{n=1}^N \sum_{m=1}^N \alpha_n \alpha_m y_n y_m x_n^T x_m - \sum_{n=1}^N \alpha_n \right)$$

$$B) \min_{\alpha} \left(\frac{1}{2} \sum_{n=1}^N \sum_{m=1}^N \alpha_n \alpha_m y_n y_m K(x_m, x_n) - \sum_{n=1}^N \alpha_n \right) \quad K(x, z) = e \left(\frac{-\|x - z\|^2}{2\sigma^2} \right)$$

Sabe-se que foram impostas as seguintes restrições durante a minimização:

$$0 \leq \alpha_n \leq 10, \quad n = 1, \dots, N \quad \sum_{n=1}^N \alpha_n y_n = 0$$

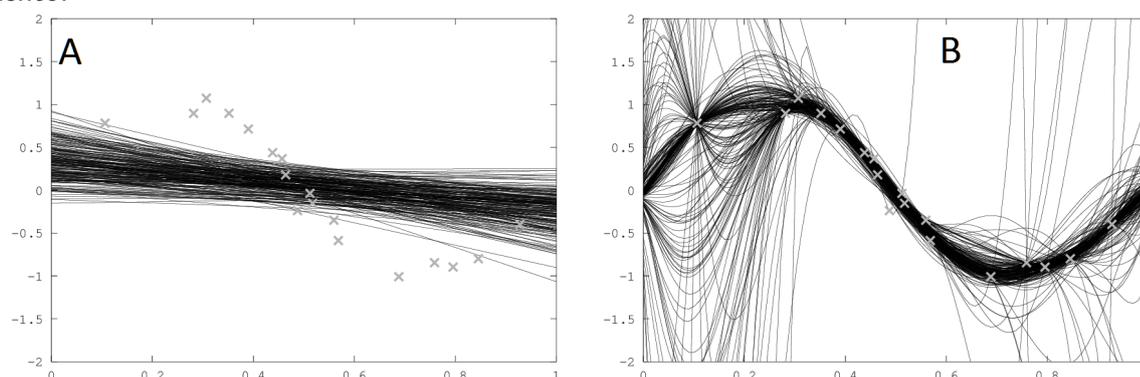
Infelizmente, não foi registado qual dos dois modelos foi usado para classificar o conjunto.

4.a) Na imagem na sua folha de resposta a linha mais grossa indica a superfície de decisão e as linhas mais finas as margens. Assinale os pontos correspondentes a valores de α maiores que zero (trace um círculo à volta dos pontos que seleccionar).

4.b) Identifique, justificando, qual dos modelos (A ou B) descritos acima foi usado para classificar este conjunto, resultando na superfície de decisão representada na imagem.

4.c) Nas condições descritas acima, eventualmente com outro conjunto de treino, seria possível obter um classificador que classificasse incorrectamente algum ponto do conjunto de treino? Justifique a sua resposta.

Pergunta 5 [2 valores] As figuras abaixo mostram o resultado do treino de dois modelos de regressão (linhas pretas) em 200 réplicas obtidas por bootstrapping a partir dos pontos indicados a cinzento.



Indique, justificando, qual dos dois modelos, A ou B, escolheria para criar uma função de regressão por bagging (a média das 200 funções de regressão obtidas treinando o modelo nas réplicas do conjunto de treino).

AA - Teste 1 - 2014-04-30

Numero: _____

0	<input type="checkbox"/>				
1	<input type="checkbox"/>				
2	<input type="checkbox"/>				
3	<input type="checkbox"/>				
4	<input type="checkbox"/>				
5	<input type="checkbox"/>				
6	<input type="checkbox"/>				
7	<input type="checkbox"/>				
8	<input type="checkbox"/>				
9	<input type="checkbox"/>				

Preencha o seu nome abaixo e o seu número à direita. Pinte por baixo de cada dígito do seu número o círculo correspondente. Por fim indique o número de alunos à sua frente e à sua direita pintando o círculo correspondente nos números abaixo.

Nome:

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Alunos à Frente	<input type="checkbox"/>															
Alunos à Direita	<input type="checkbox"/>															

1a)

1b)

1c)

2a)

2b)

2c)

AA - Teste 1 - 2014-04-30

Numero: _____

Preencha o seu nome abaixo e o seu número à direita. Pinte por baixo de cada dígito do seu número o círculo correspondente. Por fim indique o número de alunos à sua frente e à sua direita pintando o círculo correspondente nos números abaixo.

Nome:

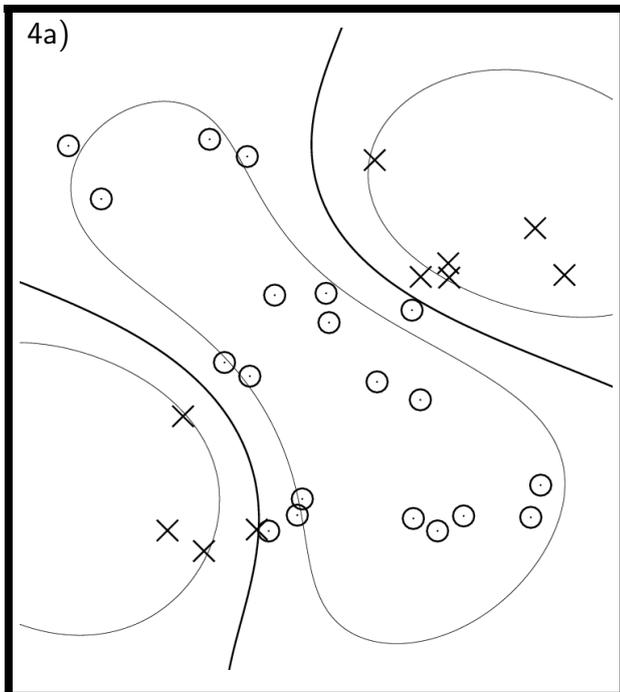
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Alunos à Frente	<input type="radio"/>															
Alunos à Direita	<input type="radio"/>															

0	<input type="radio"/>				
1	<input type="radio"/>				
2	<input type="radio"/>				
3	<input type="radio"/>				
4	<input type="radio"/>				
5	<input type="radio"/>				
6	<input type="radio"/>				
7	<input type="radio"/>				
8	<input type="radio"/>				
9	<input type="radio"/>				

3a)

3b)

3c)



4b)

4c)

5