

1º Teste de Aprendizagem Automática

2 páginas com 9 perguntas mais 2 folhas de resposta. Duração: 1 hora e 30 minutos
DI, FCT/UNL, 25 de Outubro de 2017

Pergunta 1 [1.5 valores] Explique porque é que se costuma minimizar o erro quadrático médio para treinar modelos em problemas de regressão.

Pergunta 2 [2.5 valores]
Separou-se, aleatoriamente, um conjunto de 150 pontos bidimensionais (x e y) em 3 conjuntos de 50 pontos cada (conjuntos A, B e C). A tabela à direita mostra o erro quadrático médio da previsão do y em função do x , para cada conjunto, usando polinómios de grau 2 a 7 cujos coeficientes foram obtidos minimizando o erro quadrático médio medido no conjunto A. Indique o melhor grau para o polinómio a usar nesta regressão e indique uma estimativa não tendenciosa do erro verdadeiro dessa hipótese. Se houver alternativas, indique todas. Justifique a sua resposta.

Grau	Conj. A	Conj B	Conj C
2	3.8	4.2	3.9
3	3.1	3.5	3.3
4	1.2	1.6	2.1
5	0.9	1.9	2.0
6	0.6	6.2	7.7
7	0.1	6.6	9.6

Pergunta 3 [2 valores] O erro de validação cruzada dá uma estimativa do erro do modelo ou da hipótese? Justifique a sua resposta.

Pergunta 4 [2 valores] Explique o que é a regularização e descreva um exemplo de como se usa a regularização num destes casos (à sua escolha): regressão polinomial, regressão logística ou redes neuronais.

Pergunta 5 [1 valores] Indique como desenhar uma rede neuronal de forma a poder separar classes de exemplos que não sejam linearmente separáveis e explique como é que isso permite tal separação.

Pergunta 6 [4 valores] Para criar um classificador, estimou-se a probabilidade conjunta da classe e dos valores dos atributos multiplicando a probabilidade a priori de um ponto pertencer a cada classe (a proporção da classe no conjunto de treino) pela probabilidade condicional do valor de cada atributo nessa classe. Assim, para classificar um ponto, procura-se a classe que maximiza a probabilidade conjunta de acordo com esta expressão:

$$\operatorname{argmax}_{k \in \{0,1,\dots,K\}} \ln p(C_k) + \sum_{j=1}^N \ln p(x_j|C_k)$$

6.a) Explique o que estamos a assumir acerca das distribuições de probabilidades dos atributos para que a expressão da probabilidade conjunta seja assim tão simples.

6.b) Se usarmos Kernel Density Estimation para estimar os logaritmos das probabilidades condicionais dos atributos em cada classe ($\ln p(x_j|C_k)$) podemos considerar este classificador um exemplo de *lazy learning*? Justifique a sua resposta.

Pergunta 7 [4 valores]

A figura à direita mostra os dados, a linha discriminante entre as duas classes (linha contínua) e as margens (traçado) para um classificador que prevê a classe de cada exemplo \vec{x}_t pelo sinal da seguinte expressão:

$$\sum_{n=1}^N \alpha_n y_n K(\vec{x}_n, \vec{x}_t)$$

Os valores de α_n para os vectores do conjunto de treino foram calculados pela solução desta minimização:

$$\min_{\alpha} \left(\frac{1}{2} \sum_{n=1}^N \sum_{m=1}^N \alpha_n \alpha_m y_n y_m K(\vec{x}_m, \vec{x}_n) - \sum_{n=1}^N \alpha_n \right)$$

sujeita às seguintes restrições:

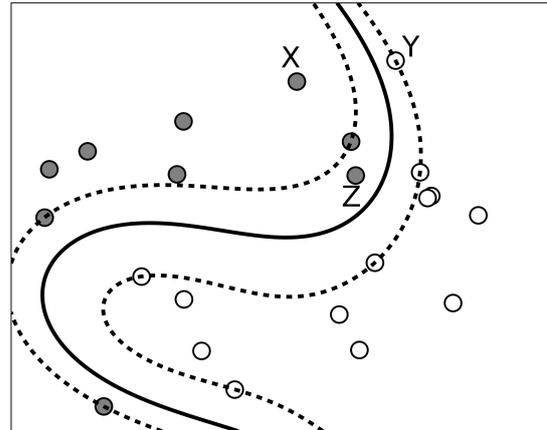
$$0 \leq \alpha_n \leq 5, \quad n = 1, \dots, N \qquad \sum_{n=1}^N \alpha_n y_n = 0$$

A função $K(\vec{x}_m, \vec{x}_n)$ foi uma destas duas:

- Opção A: $K(\vec{x}_m, \vec{x}_n) = \vec{x}_m^T \vec{x}_n$
- Opção B: $K(\vec{x}_m, \vec{x}_n) = (\vec{x}_m^T \vec{x}_n)^3$

7.a) Qual das duas opções (A ou B) foi usada para a função $K(\vec{x}_m, \vec{x}_n)$? Justifique a sua resposta.

7.b) Indique os valores dos multiplicadores α para os vectores X, Y e Z indicados na imagem. Se não puder determinar o valor exacto do α em algum destes casos, indique o intervalo de valores possíveis para esse vector.



Pergunta 8 [1.5 valores] Explique porque é que não é útil usar o método de Bootstrap Aggregating (Bagging) com modelos cujo *bias* seja maior do que a variância.

Pergunta 9 [1.5 valores] No algoritmo AdaBoost, treina-se sempre o mesmo modelo a cada iteração, com os mesmos dados. No entanto, a cada iteração obtém-se hipóteses diferentes. Explique porquê.

AA, Teste 1. 2017-10-25

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 filas de alunos à sua frente e o número de alunos à sua direita pintando o círculo correspondente abaixo.

Nome:

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Filas 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"/>				

1)

2)

3)

4)

5)

AA, Teste 1. 2017-10-25

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 filas de alunos à sua frente e o número de alunos à sua direita pintando o círculo correspondente abaixo.

Nome:

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Filas 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"/>				

6 a)

6 b)

7 a)

7 b)

8)

9)