

1º Teste de Aprendizagem Automática

2 pages with 9 questions plus 2 answer sheets. Duration of 1 hour and 30 minutes
DI, FCT/UNL, 25 October 2017

Pergunta 1 [1.5 valores] Explain why one usually minimizes the mean quadratic error to train models in regression problems.

Pergunta 2 [2.5 valores]

We randomly split a set of 150 bidimensional points (x and y) in 3 sets of 50 points each (sets A, B and C). The table on the right shows the mean quadratic error for the prediction of y as a function of x , for each set, using polynomial curves of degree 2 through 7 whose coefficients were obtained minimizing the mean quadratic error to the points in set A.

Indicate the best degree for the polynomial to use in this regression and an unbiased estimate of the true error for the resulting hypothesis. If there are alternatives, indicate all of them. Justify your answer.

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] Cross-validation error gives us an estimate of the error of the model or of the hypothesis? Justify your answer.

Pergunta 4 [2 valores] Explain what regularization is and describe how to use regularization in one of these cases (choose one): polynomial regression, logistic regression or neural networks.

Pergunta 5 [1 valores] Explain how to design a neural network so that it can separate classes of examples that are not linearly separable and how such design allows non-linear separation.

Pergunta 6 [4 valores] To create a classifier, we estimated the joint probability of the class and attribute values by multiplying the a priori probability of a point belonging to a class (the proportion of the class in the training set) by the conditional probability of each attribute value given that class. Thus, to classify a point, we find the class that maximizes the joint probability according to this expression:

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

6.a) Explain what we are assuming about the probability distributions of the attributes so that the expression of the joint probability is as simple as this one.

6.b) If we use Kernel Density Estimation to estimate the logarithms of the conditional probabilities of the attributes in each class ($\ln p(x_j|C_k)$) can we consider this classifier to be an example of lazy learning? Justify your answer.

Pergunta 7 [4 valores]

The figure on the right shows the data, the discriminant line (solid line) and the margins (dashed line) for a classifier that predicts the class of each example \vec{x}_t according to the sign of the following expression:

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

The α_n values for the vectors in the training set were computed by solving the following minimization:

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

subject to the following restrictions:

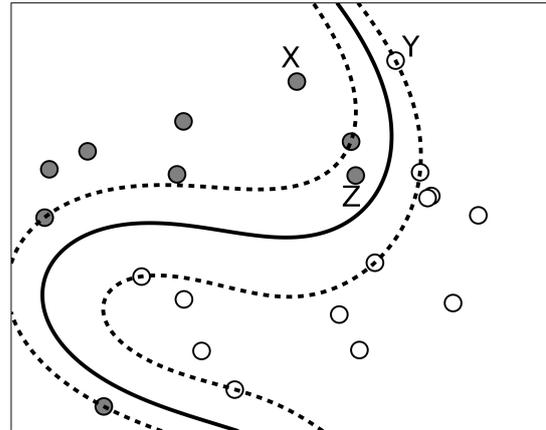
$$0 \leq \alpha_n \leq 5, \quad n = 1, \dots, N \qquad \sum_{n=1}^N \alpha_n y_n = 0$$

The function $K(\vec{x}_m, \vec{x}_n)$ was one of the following two:

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

7.a) Which of the two options (A or B) was used for $K(\vec{x}_m, \vec{x}_n)$? Justify your answer.

7.b) Write the values for the multipliers α for vectors X, Y e Z shown on the image. If you cannot tell the exact value of α in any of these cases, indicate the interval of possible values for that vector.



Pergunta 8 [1.5 valores] Explain why it is not useful to use the Bootstrap Aggregating (Bagging) on models whose *bias* is greater than their variance.

Pergunta 9 [1.5 valores] In the AdaBoost algorithm, we always train the same model at each iteration. However, at each iteration we get different hypotheses. Explain why.

AA, Teste 1. 2017-10-25

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 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="checkbox"/>															
Alunos à Direita	<input type="checkbox"/>															

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)