

2º Teste de Aprendizagem Automática

3 páginas com 8 perguntas e 2 folhas de resposta. Duração: 2 horas
DI, FCT/UNL, 10 de Dezembro de 2015

Pergunta 1 [2 valores] A seguinte expressão indica o limite superior provável, com probabilidade $1 - \delta$, para o erro verdadeiro de uma hipótese \hat{h} obtida por minimização do erro empírico, em função da soma de dois termos. O primeiro termo é o menor erro verdadeiro possível no espaço finito de hipóteses \mathcal{H} . O segundo termo é função do número total de hipóteses em \mathcal{H} , o tamanho do conjunto de treino, m , e a probabilidade δ .

$$E(\hat{h}) = \left(\arg \min_{h \in \mathcal{H}} E(h) \right) + 2\sqrt{\frac{1}{2m} \ln \frac{|\mathcal{H}|}{\delta}}$$

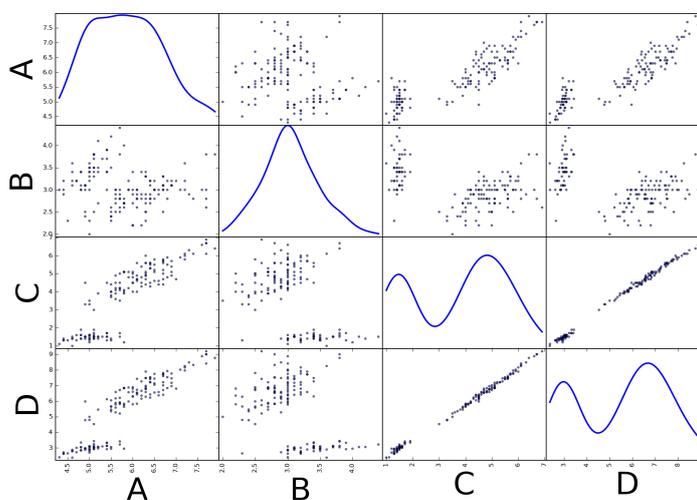
Considere também a seguinte frase: “Treinando dois modelos de classificação de forma a minimizar o erro no conjunto de treino, aquele que apresentar o menor erro no conjunto de treino terá provavelmente também o menor erro verdadeiro.”

1.a) Indique que termo da expressão acima tem de ser dominante para que a frase seja **verdadeira**, e **justifique a sua resposta**.

1.b) Indique que termo da expressão acima tem de ser dominante para que a frase seja **falsa**, e **justifique a sua resposta**.

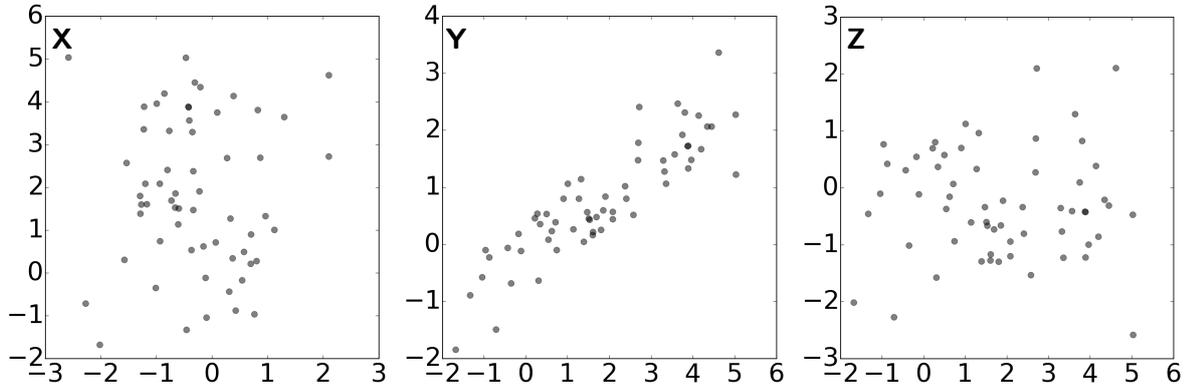
Pergunta 2 [2 valores] Em geral, nesta disciplina temos treinado classificadores minimizando o erro de classificação no conjunto de treino. Em que situações é que minimizar o erro de classificação durante o treino não é a melhor abordagem? Justifique a sua resposta e indique o que se deve fazer nesses casos.

Pergunta 3 [2 valores] A figura mostra a matriz de plots de todos os pares de atributos de um conjunto de dados com quatro atributos, A, B, C e D. Os gráficos na diagonal são Kernel Density Estimations das distribuições de cada atributo. Sabendo que este conjunto de dados não está etiquetado e será usado apenas para aprendizagem não supervisionada, e querendo filtrar um dos atributos para reduzir o conjunto a três atributos, indique, justificando, qual dos atributos eliminaria.

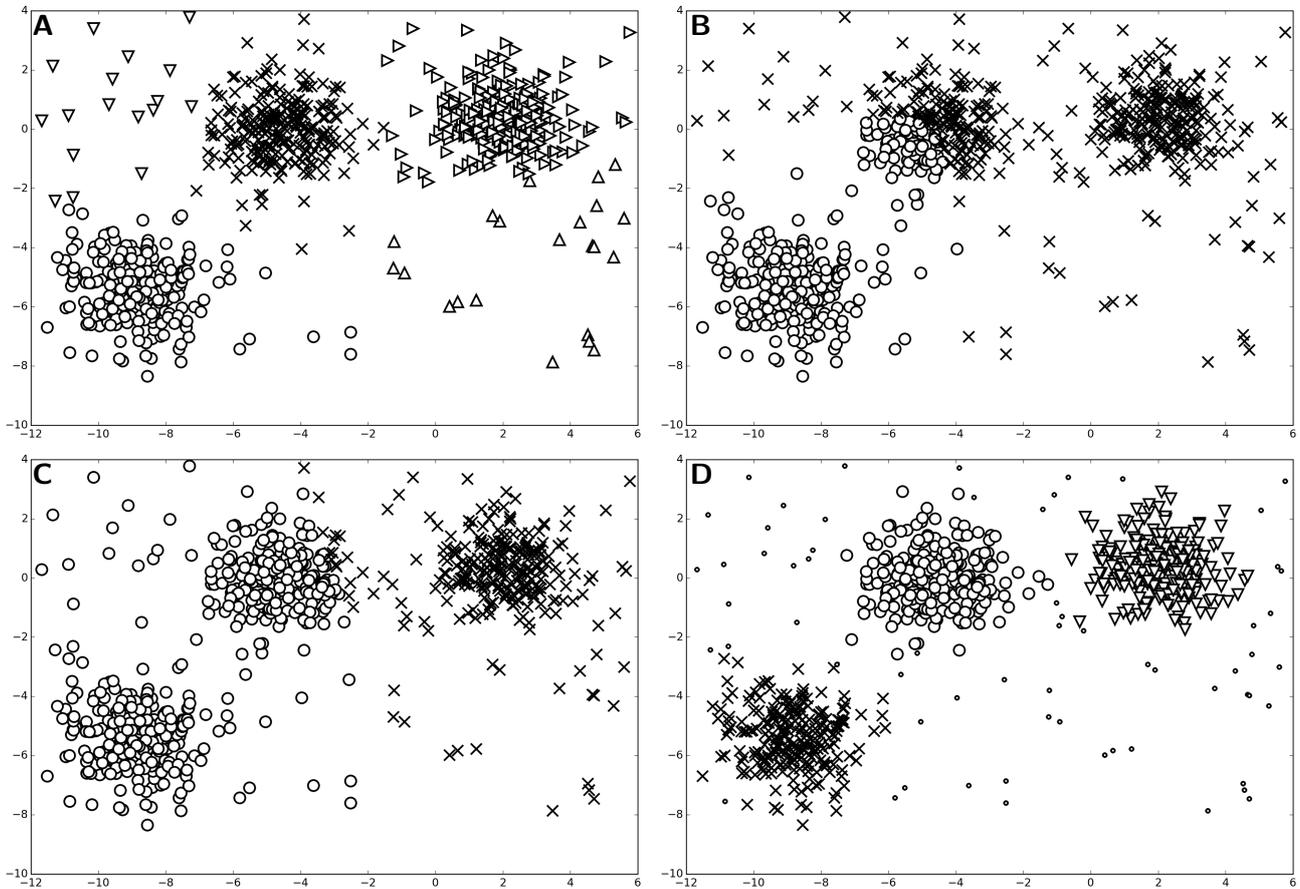


Pergunta 4 [2 valores] Descreva o algoritmo de treino de um mapa auto-organizável (*self-organizing map*, SOM). Assuma que eu já conheço a arquitectura desta rede, focando apenas o treino.

Pergunta 5 [2 valores] Para reduzir a dimensionalidade de um conjunto de dados foi feita uma análise de componente principal (Principal Component Analysis). O conjunto de dados foi projectado nos dois vectores próprios da matriz de covariância que tinham os maiores valores próprios. Aquele que tinha o maior valor próprio foi usado para traçar a coordenada x (horizontal) e o outro, com o segundo maior valor próprio, foi usado a coordenada y (vertical). Indique, **justificando**, qual dos três gráficos X, Y ou Z, representa o resultado desta projecção.



Pergunta 6 [6 valores] O mesmo conjunto de dados foi agrupado com quatro algoritmos de *clustering* diferentes. Leia as quatro alíneas desta pergunta primeiro, consultando o gráfico que se segue, e depois responda a cada uma **justificando a sua resposta**. No gráfico, cada símbolo maior representa um ponto num *cluster*. Os pontos que não foram atribuídos a um *cluster* são representados com círculos pequenos (só ocorre no gráfico D).



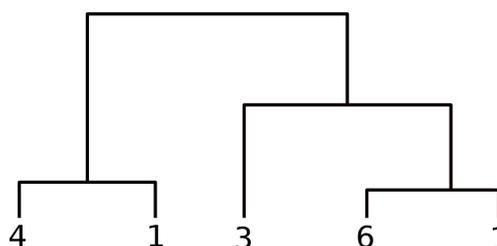
6.a) Qual dos quatro gráficos (A, B, C, D) mostra o resultado da aplicação do algoritmo **Density-based spatial clustering of applications with noise (DBSCAN)**, com um valor de 10 para o número mínimo de vizinhos para os pontos de *core*?

6.b) Qual dos quatro gráficos (A, B, C, D) mostra o resultado de *clustering* usando uma **mistura de duas distribuições Gaussianas**?

6.c) Qual dos quatro gráficos (A, B, C, D) mostra o resultado de *clustering* usando o algoritmo de **k-means** com um valor de $k = 2$?

6.d) Qual dos quatro gráficos (A, B, C, D) mostra o resultado de *clustering* usando o algoritmo de **Affinity Propagation** com um valor inicial de propensão (a diagonal da matriz de similaridade) igual a -200?

Pergunta 7 [2 valores] O gráfico à direita representa o dendrograma completo produzido por um algoritmo de *clustering* hierárquico. Os números nas folhas do dendrograma indicam o número de elementos do conjunto de treino que ficaram no *cluster* correspondente a cada folha. Indique, justificando, se este é o resultado de um algoritmo de *clustering* aglomerativo (*agglomerative clustering*) ou divisivo (*divisive clustering*)



Pergunta 8 [2 valores] Responda a **uma** das duas perguntas abaixo:

- Num modelo oculto de Markov, a variável oculta pode adoptar um de N estados e emitir um de K símbolos em cada estado. Quantos e quais valores é preciso especificar para instanciar completamente este modelo?
- Quando uma rede neuronal artificial tem muitas camadas, o algoritmo de *backpropagation* torna-se demasiado ineficiente e instável para treinar a rede a partir de um conjunto inicial de pesos gerados aleatoriamente. Descreva uma forma usada em *deep learning* para contornar este problema.

AA, Teste 2, 2015-12-10

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"/>				

1a)

1b)

2)

3)

4)

5)

AA, Teste 2, 2015-12-10

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"/>				

6a) DBSCAN

6b) Mistura de Gaussianas

6c) k-means

6d) Affinity Propagation

7)

8)