



# MO 906 - Introdução à Inteligência Artificial

## 2º Semestre de 2009

### Lista 2

#### 1 Redes Bayesianas.

1. “Russell & Norvig” ex. 14.1 (página 517).
2. “Russell & Norvig” ex. 14.3 (página 518).
3. “Russell & Norvig” ex. 14.4 (página 519).
4. O que é equivalência de Markov? Escreva um algoritmo em tempo polinomial que determine se duas DAGs são equivalentes segundo o critério de Markov (“Markov-equivalent”).
5. “Neapolitan” ex 2.17 (página 118).
6. “Neapolitan” ex 2.28 (página 121).
7. “Neapolitan” ex 3.1 (página 181).
8. “Neapolitan” ex 3.2 (página 181).
9. “Neapolitan” ex 6.1 (página 367).
10. “Neapolitan” ex 6.2 (página 367).
11. “Neapolitan” ex 6.4 (página 367).

#### 2 Classificadores e Aprendizado.

1. Como pode se reduzir o problema de classificação em duas classes ao de representação de funções? Como se faz o mesmo para classificação em  $C$  classes?
2. Defina conjunto de treinamento e conjunto de teste. Como eles se relacionam ao problema de “overfitting”?
3. O que é validação cruzada e para que serve? Compare a técnica de “leave-one-out” com validação cruzada.
4. Qual a interpretação do método de mínimos quadrados para regressão linear com classificadores?
5. “Russell & Norvig” ex. 20.21 (página 737).

### 3 K-Vizinhos, K-Means.

1. Leia “Russell & Norvig” 20.3 (703-705).
2. Leia “Russell & Norvig” 20.4 (710-713).
3. Proponha um método para acelerar a busca dos K-Vizinhos mais próximos - analise a complexidade da sua proposta.

### 4 Redes Neurais.

1. “Mitchell” ex. 4.1 (página 124).
2. “Mitchell” ex. 4.2 (página 124).
3. “Mitchell” ex. 4.3 (página 124).
4. “Mitchell” ex. 4.4 (página 124).
5. “Mitchell” ex. 4.5 (página 124).
6. “Mitchell” ex. 4.8 (página 125).
7. “Mitchell” ex. 4.9 (página 125).
8. “Mitchell” ex. 4.10 (página 125).

### 5 Redução de Dimensionalidade.

1. Explique o que é PCA e LDA.
2. Mostre como resolver o problema dos autovalores generalizados:  $\max a^\top B a$  sujeito a  $a^\top W a = 1$  através de uma redução a um problema de autovalores comum.
3. Dado o seguinte conjunto de dados:  $\{(1,2,1), (3,-2,7), (8,0,-2), (-3,3,1), (-1,-1,-1), (5,5,5), (4,3,4), (-2,0,2), (0,1,0), (-1,-1,1), (2,2,-3), (4,-1,-1)\}$ , faça a decomposição em componentes principais (PCA), i.e. calcule os três componentes principais, e projete os dados neste novo espaço. Como uma segunda etapa, faça a projeção no espaço normalizado (de forma que a variância ao longo de cada componente principal seja igual à 1).
4. Dado o seguinte conjunto de dados:  
Para classe  $A = \{(1,2,1), (3,-2,7), (8,0,-2), (-3,3,1), (-1,-1,-1), (5,5,5)\}$   
Para classe  $B = \{(4,3,4), (-2,0,2), (0,1,0), (-1,-1,1), (2,2,-3), (4,-1,-1)\}$ .  
Faça a LDA deste conjunto, e encontre a projeção em  $\mathbb{R}$  de cada um dos elementos do conjunto de dados.
5. Dado um conjunto de dados em  $\mathbb{R}^n$ , aplica-se uma decomposição PCA e utiliza-se todos os  $n$  componentes principais para o novo conjunto de dados. Compare a performance de um classificador linear nos dois casos.
6. Considere um espaço tridimensional, com coordenadas  $x$ ,  $y$  e  $z$ , no qual existe uma determinada distribuição de padrões. Nesta distribuição,  $z$  obedece à equação  $z = 2x + y$ , e  $x$  e  $y$  são independentes entre si, com as seguintes médias e variâncias  $\mu_x = 0$ ,  $\sigma_x^2 = 1$ ,  $\mu_y = 0$ ,  $\sigma_y^2 = 3$ . Determine a primeira direção principal desta distribuição (indique essa direção por meio de um vetor).

7. Considere um conjunto de dados bidimensionais obtidos de uma Gaussiana Multidimensional com média igual a zero. A análise por componentes principais destes dados resulta nos dois autovetores  $(\sqrt{2}/2, \sqrt{2}/2)$  e  $(\sqrt{2}/2, -\sqrt{2}/2)$  com autovalores 2 e 3 respectivamente.
- (a) Determine as variâncias  $\mu_1$  e  $\mu_2$  ao longo dos dois eixos do espaço dos dados originais.
  - (b) Indique a componente principal com maior energia. Qual é a sua variância como uma percentagem da variância total dos dados?
  - (c) Determine o valor dos pontos  $(3, 2)$  e  $(3, -2)$  quando projetados na componente principal do ítem anterior.
  - (d) Reconstrua para o espaço original os pontos projetados do ítem anterior. Calcule o valor absoluto do erro de reconstrução para cada um dos dois pontos. Explique de forma geométrica porque um possui erro menor do que o outro.

## 6 Coletâneas de Classificadores.

1. Compare as estruturas de coletâneas em paralelo, em cascata e hierárquicas. Quais as suas vantagens e desvantagens?
2. Descreva a técnica de Bagging.
3. Descreva a técnica de Boosting.