



MO 906 - Introdução à Inteligência Artificial

1º Semestre de 2005

Lista 3

Sem Entrega

1 Probabilidade.

- Uma caixa com N lâmpadas contém r lâmpadas ($r < N$) com filamento partido.
 - Essas lâmpadas são verificadas uma a uma, até que uma lâmpada defeituosa seja encontrada. Descreva um espaço amostral para este experimento.
 - Suponha que as lâmpadas acima sejam verificadas uma a uma, até que todas as defeituosas tenham sido encontradas. Descreva o espaço amostral para este experimento.
- Um lote contém peças pesando 5, 10, 15, \dots , 50 gramas. Admitamos que ao menos duas peças de cada peso sejam encontradas no lote. Duas peças são retiradas do lote. Sejam X o peso da primeira peça escolhida e Y o peso da segunda. Portanto, o par de números (X, Y) representa um resultado simples do experimento. Empregando o plano XY , marque o espaço amostral e os seguintes eventos:
 - $\{X = Y\}$.
 - $\{Y > X\}$.
 - A segunda peça é duas vezes mais pesada que a primeira.
 - A primeira peça pesa menos 10 gramas que a segunda peça.
 - O peso médio de duas peças é menor do que 30 gramas.
- Durante um período de 24 horas, em algum momento X , uma chave é posta na posição “ligada”. Depois, em algum momento futuro Y (ainda durante o mesmo período de 24 horas), a chave é virada para a posição “desligada”. Suponha que X e Y sejam medidas em horas, no eixo dos tempos, com o início do período na origem da escala. O resultado do experimento é constituído pelo par de números (X, Y) .
 - Descreva o espaço amostral.
 - Descreva e marque no plano XY os seguintes eventos:
 - O circuito está ligado por uma hora ou menos.

- ii. O circuito está ligado no tempo z , onde z é algum instante no período dado de 24h.
 - iii. O circuito é ligado antes do tempo t_1 e desligado depois do tempo t_2 (onde também $t_1 < t_2$ são dois instantes durante o período de 24 horas especificado).
 - iv. O circuito permanece ligado duas vezes mais tempo do que desligado.
4. Sejam A , B e C três eventos associados a um experimento. Exprima, em notações de conjuntos, as seguintes afirmações verbais:
- (a) Ao menos um dos eventos ocorre.
 - (b) Exatamente um dos eventos ocorre.
 - (c) Exatamente dois eventos ocorrem.
 - (d) Não mais de dois dos eventos ocorrem simultaneamente.

2 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).

3 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. Qual a interpretação do método de mínimos quadrados para regressão linear com classificadores?
4. “Russell & Norvig” ex. 20.21 (página 737).

4 K-Vizinhos, K-Means.

1. Leia “Russell & Norvig” 20.3 (703-705).
2. Leia “Russell & Norvig” 20.4 (710-713).

5 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).

6 Redução de Dimensionalidade.

1. Explique o que é PCA e LDA.
2. Mostre como resolver o problema dos autovalores generalizados: $\max a^T B a$ sujeito a $a^T W a = 1$ 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.

7 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.