Novas Tecnologias de Informação
Licenciatura em Estatística e Gestão de Informação
Cotação: Grupo I - 1 valor cada; Grupo II-1,4,1,1 Grupo III-1,2
ATENÇÃO: Cada pergunta de escolha múltipla errada desconta 0.4 valores
Duração: 2 horas Teste A

Doutor Moura-Pires & Doutor Victor Lobo **Ano lectivo: 2002-2003 – 2ª CHMADA**

I

Escolha uma e uma só resposta para cada uma das seguintes questões

I.1) Imagine que um colega seu lhe perguntava se as "redes neuronais artificias" podiam, por exemplo, fazer contas de multiplicar. A resposta certa seria:

a) Sim: podem fazê-lo com várias vantagens sobre os métodos tradicionais, nomeadamente sendo mais rápidas a fazer as operações.
b) Não: podem fazer outras operações, mas como os neurónios implementam funções lineares são incapazes de fazer multiplicações.
c) Sim: no entanto, havendo um algoritmo determinístico eficiente e exacto, é um desperdício perder tempo a treinar uma rede para essa tarefa, e ainda por cima não haveria a garantia que o resultado fosse sempre correcto.
d) Não: uma rede neuronal serve para fazer classificações mas não para fazer operações aritméticas.

I.2) Ao longo desta cadeira foram dados várias técnicas para aprender, a partir de dados, a fazer classificações. Qual das afirmações seguintes é falsa:

a) Os métodos heurísticos que foram dados não garantem que o classificador final seja óptimo (no sentido em que tem a menor taxa de erro possível).
b) Existe um limite à exactidão que é possível obter num dado problema, que é a exactidão do classificador de Bayes.
c) Nenhuma das heurísticas é melhor que qualquer das outras em todos os problemas.
d) As redes neuronais conseguem obter taxas de erro menores que os outros métodos heurísticos.

I.3) Qual das afirmações seguintes é verdadeira:

a) As fronteiras entre classes definidas por uma árvore de decisão podem ser sigmoides

b) Uma das vantagens das árovres de decisão é que são sempre mais simples (no sentido que exigem menos operações para chegar a um resultado) do que uma rede neuronal.
c) Uma das vantagens das árovres de decisão sobre outros métodos de classificação é que geralmente podem dar explicações para as suas classificações que são facilmente entendíveis por humanos.
d) As árvores de decisão, no conteixto em que foram dadas nesta cadeira, e ao contrário das redes neuronais, não necessitam de um conjunto de treino para serem construídas.

I.4) Uma cadeia de hipermercados pretende abrir lojas mais pequenas orientadas para sectores de mercado mais restritas. Como base para essa segmentação de mercado, dispõe dos registos das máquinas registadores referents às compras efectuadas pelos seus clientes. Aconselharia a utilização de:

a) Uma rede neuronal multicamada, treinada com backprogation, em que usaria como variáveis de entrada o montante das compras que os clientes efectuam em cada tipo de produtos, e como variável de saída o segmento pretendido
b) Uma rede neuronal SOM, em que usaria como variáveis de entrada o montante das compras que os clientes efectuam em cada tipo de produtos.
c) Uma árvore de decisão, em que usaria como variáveis de entrada o montante das compras que os clientes efectuam em cada tipo de produtos, e como variável de saída o segmento pretendido.
d) Um algoritmo genético, em que usaria como variáveis de entrada o montante das compras que os clientes efectuam em cada tipo de produtos.

I.5) Qual das seguintes afirmações é falsa ?

a) Para que se possam aplicar métodos de aprendizagem automática é necessário dispôr de um conjunto de dados de treino para os quais quer as variáveis de entrada quer as classes (ou valores de saída no caso de regressão) sejam conhecidos.
b) É possível aplicar métodos de aprendizagem automática mesmo quando não se conhece a classe "verdadeira" dos dados usados para treino.
c) Num problema de aprendizagem não supervisionada não é possível calcular taxas de erro.
d) Há alguns métodos de aprendizagem automática que não usam conjuntos de validação.

I.6) Para resolver o problema clássico do caixeiro viajante (discutido nas aulas) pode-se:

a) Usar um SOM
b) Usar um algoritmo genético
c) Usar uma pesquisa tabu
d) Usar qualquer das 3 abordagens (SOM, algoritmos genétocos, pesquisas tabu)

I.7) Qual das seguintes afirmações é falsa ?

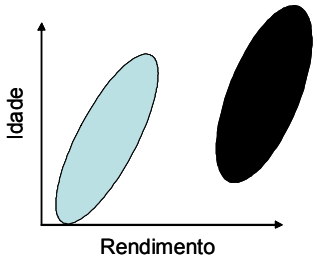
Uma rede RBF tende a ser melhor para aprendizagem local, enquanto uma rede multicamada tende a generalizar melhor pera dados que difiram mais do conjunto de treino.
Uma rede neuronal SOM pode ser adaptada para problemas de classificação.
O algoritmo de backpropagation (dado nas aulas) é o método mais rápido que garante a congergência da rede para solução óptima.
Uma rede neuronal multicamada pode ser usada para problemas de previsão.

I.8) O critério de Gini é por vezes usado para escolher de entre as diversas partições possíveis para árvores de decisão porque:

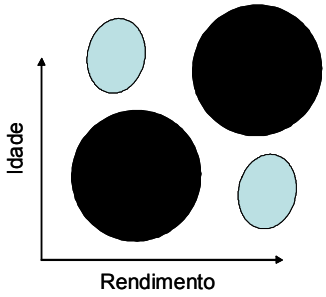
a) É aquele que geralmente produz resultados mais exactos
b) Tem mais significado do que a utilização da entropia
c) É mais fácil de calcular que a maioria dos outros critérios
d) Não exige o cáclulo de probabilidades

Cada um dos métodos de previsão estudado tem vantagens e desvantagens que dependem muito dos dados em questão. Normalmente os dados têm muitos atributos, tornando a sua visualização difícil. No entanto, se os dados forem bi-dimensionais, a visualização é trivial, e por vezes pode ser óbvio escolher o melhor classificador, quer em termos de taxa de erro quer em termos de facilidade de implementação. Nas questões seguintes, são apresentados gráficos com diversas distribuições de dados. Os dados (inventados) representam clientes de um banco, descritos através do seu rendimento bruto e da sua idade. Um perito dividiu os clientes que provavelmente irão aderir a uma campanha do banco (representados a escuro) e os que em princípio não estarão interessados (representados a claro). Para cada uma das situações indique qual o tipo de classificador mais apropriado (i.e., mais simples, e com menor taxa de erro)..

I.9)
a) Perceptrão simples
b) Perceptrão multicamada
c) Árvore de decisão
d) Rede neuronal RBF



I.10)
a) Perceptrão simples
b) Perceptrão multicamada
c) Árvore de decisão
d) Rede neuronal RBF



II

Com a chegada do verão, há sempre um aumento dos acidentes nas praias. Verifica-se que quanto mais quente fôr o dia, maior a probabilidade de haver acidentes. Verifica-se também que quanto pior fôr o estado de mar (medido pela altura da vaga), mas acidentes há. No entanto, a Marinha (mais concretamente o Instituto de Socorros a Náufragos), gostaria de prever com o maior rigôr possível o número de acidentes, de modo a poder reforçar convinientemente a vigilância. Para tal, gostaria de ter um sistema que dada a previsão da temperatura máxima e da ondulação estimasse o número de acidentes. Como a relação entre essas variáveis é relativamente simples, prevê-se que um único perceptrão, com uma função de activação linear (com declive 1), e treinado com algorimo de backpropagation (neste caso simplificado ao máximo dado tratar-se de uma única camada) resolva o problema. Para treinar essa rede dispõe dos dados relativos aos acidentes numa dada zona da costa duante o ano passado . Desses dados, foram seleccionados os 9 apresentados na tabla ao lado para conjunto de treino. A rede pode ser inicializada com quaisquer valores, mas por hipótese foi seleccionado o valor 3 para todos os pesos sinápticos, e um "bias" de 70 (assuma que esse bias é multiplicado por um termo constante de -1 antes de ser somado no neurónio).

Temperatura	Ondulação	Acidentes
24	0,5	0
23	2,0	12
36	1,6	31
38	1,0	30
25	1,1	7
29	0,8	11
33	1,8	28
30	0,6	11
35	0,1	16

II.1) Se usasse o software SAS Enterprise Miner para resolver este problema, que valor tentaria minimizar para treinar a rede ?

II.2) Como não dispõe do SAS Enterprise Miner durante o teste, faça o treino da rede "à mão", apresentando os cálculos. Para simplificar, use um ritmo de aprendizagem constante e igual a 0,4, não utilize momentos, e faça apenas 3 iterações, isto é, actualize os pesos da rede apenas 3 vezes usando os 3 primeiros dias da tabela.

II.3) Qual seria o erro da rede treinada na alínea anterior (se não fez essa alínea use a rede original), para o caso do último dia da tabela ?

III.4) Treinando a rede até atingir um erro 0 no conjunto de treino, obtém um peso sináptico de 1,8 para a temperatura, 9,1 para a ondulação, e um bias de 48. Usando esta rede, quantos acidentes prevê para um dia em que a temperatura suba aos 40 graus, e a ondulação seja 0 (mar estanhado) ?

III

III.1) Explique qual a diferença entre um classificador Bayesiano MAP (maximum a posteriori) e um classificador naive de Bayes. Explique em que casos se deve optar por um ou outro desses dois tipos de classificadores, e porquê.

III.2) Imagine que quer classificar as pessoas em 3 grupos (desportista, normal, sedentário), basado no seu peso (dado em intervalos de 5 Kg). Para tal dispõe de uma base de dados de 500 pessoas, que foram atribuídas a esses 3 grupos. Sabe-se que há mais pessoas normais do que desportistas ou sedentários. Entre os classificadores de ML (máxima verosimilhança), MAP e naive de Bayes dados nas aulas, qual deles escolheria, e porquê.

Boa sorte !