S228/406

DUBLIN INSTITUTE OF TECHNOLOGY
KEVIN STREET, DUBLIN 8

BSc (Hons) in Computer Science

Stage 4

SEMESTER 2 EXAMINATIONS 2008

ARTIFICIAL INTELLIGENCE 2

Dr. John Kelleher
Prof. B. O'Shea
Dr. I. Arena

Duration: 2 Hours

Answer Question 1 (40 marks) **and**

any 2 Other Questions (30 marks each).

1. (a) Distinguish between **supervised** and **unsupervised** learning.

(5 marks)

(b) In the context of machine learning, explain what is meant by **overfitting** the training data.

(5 marks)

(c) Explain what is meant by the term **abductive reasoning**.

(5 marks)

(d) What does it mean if two classes C_1 and C_2 are described as **linearly separable**.

(5 marks)

(e) Let us say we have three classification algorithms. How can we order these three from best to worst?

(20 marks)

2. (a) In your local power station, there is an alarm that senses when a temperature gauge exceeds a given threshold. The gauge measures the temperature of the core. Consider the Boolean variables A (alarm sounds), F_A (alarm is faulty), and F_G (gauge is faulty) and multivalued nodes G (gauge reading) and T (actual core temperature).

(i) Draw a Bayesian network for this domain, given that the gauge is more likely to fail when the core temperature gets too high.

(5 marks)

(ii) Suppose there are just two possible actual and measured temperatures: normal and high. The probability that the gauge gives the correct temperature is x when it is working, but y when it is faulty. Give the conditional probability table associated with node G.

(5 marks)

(b) After your yearly checkup, the doctor has bad news and good news. The bad news is that you tested positive for a serious disease and that the test is 99% accurate (i.e., the probability of testing positive when you do have the disease is 0.99, as is the probability of testing negative when you don't have the disease). The good news is that this is a rare disease, striking only 1 in 10,000 people of your age.

(i) Why is it good news that the disease is rare?

(10 marks)

(ii) What are the chances that you actually have the disease?

(10 marks)

3. (a) In the context of Decision Tree Learning define what is meant by the following terms:

   (i) entropy

   (5 marks)

   (ii) information gain

   (5 marks)

   (b) Using the ID3 algorithm we never test the same attribute twice along one path in a decision tree. Why not?

   (5 marks)

   (c) Suppose we generate a training set from a decision tree and then apply decision-tree learning to the training-set. Is it the case that the learning algorithm will eventually return the correct tree as the training set size goes to infinity? Why or why not?

   (5 marks)

   (d) Discuss the advantages and disadvantages of k-**Nearest Neighbour** classification.

   (10 marks)

4. Figure 1 shows a backprogation network that is currently processing the training vector $[1.0, 0.9, 0.9]$ which has an associated target vector $[0.1, 0.9, 1.0]$. Given that the output from unit B is 0.6 and from C is 0.8, and assuming that the activation function used at all nodes in the network is the logistic function (i.e., $f(x) = \frac{1}{1+\exp^{-x}}$):

   (a) Calculate the actual output vector (to 3 decimal places).

   (5 marks)

   (b) Calculate the error for each output unit.

   (5 marks)

   (c) Calculate the error for each hidden unit B and C.

   (10 marks)

   (d) Calculate the new weight for the connection from unit B to the output unit D after the training example has been processed. Use a learning rate of $\alpha = 0.25$ and momentum of zero.

   (10 marks)

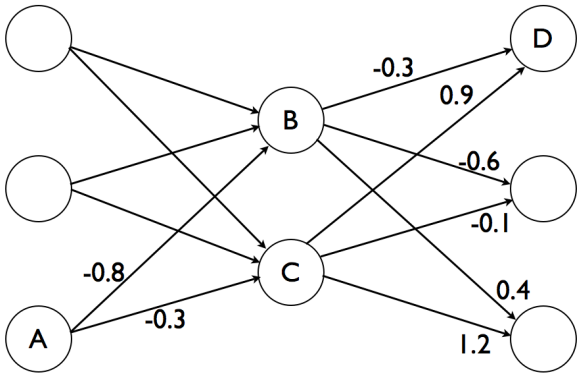

Figure 1: Example Neural Net