S228/406

DUBLIN INSTITUTE OF TECHNOLOGY
KEVIN STREET, DUBLIN 8

BSc (Hons) in Computer Science

Stage 4

SEMESTER 2 EXAMINATIONS 2010

ARTIFICIAL INTELLIGENCE 2

Dr. John Kelleher
Dr. D. Lillis
Dr. I. Arena

Duration: 2 Hours

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

any 2 Other Questions (30 marks each).

Table 1: Joint Distribution for X and Y

	$X = x_1$	$X = x_2$
$Y = y_1$	0.02	0.30
$Y = y_2$	0.14	0.32
$Y = y_3$	0.10	0.12

1. (a) Given the joint distribution for X and Y listed in Table 1 calculate:

   (i) $P(Y = y_2)$

   (5 marks)

   (ii) $P(Y = y_2|X = x_1)$

   (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. What are the chances that you actually have the disease?

   (10 marks)

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

   (20 marks)

2. (a) Consider the following scenario: *When it rains the grass usually gets wet but again not always. When the sprinkler is on the grass sometimes gets wet but not always. Also, when it rains the sprinkler is often turned of (although sometimes its left on by mistake)..*

   (i) Using the following Boolean variables $Rain$ (true of false), $Sprinker$ (on or off), and $Grass$ (wet or dry) draw a Bayesian network that models this domain.

   (5 marks)

   (ii) for each node in the network create a conditional probability table (CPT). Fill in the cells in the CPT tables with probabilities that you think are reasonable and that are not equal to 0 or 1.

   (5 marks)

   (iii) Using the Bayesian network and conditional probability tables you have created, compute the probability that it is raining give that the grass is wet?

   (5 marks)

   (b) You are on holidays on Fisher Island. The yearly weather on Fisher Island comes in five different varieties:

   - there is a 10% chance that there will be rain everyday of the year.
   - there is a 20% chance that there will be rain on 75% of the days of the year.
   - there is a 40% chance that there will be rain on 50% of the days of the year.
   - there is a 20% chance that there will be rain on 25% of the days of the year.
   - there is a 10% chance that there will be no rain on any day of the year.

   (i) given that it has rained on day 1 and 2 of the year compute the posterior probability of each of the 5 yearly weather patterns on day 2 of the year. Give your answer rounded to four places of precision.

   (5 marks)

   (ii) given that after the first 10 days of the year the weather has been such that the posterior probabilities of each of the 5 varieties of the yearly weather on Fisher Island are: there is now a 90% chance that there will be rain everyday for the rest of the year; a 7% chance that there will be rain on 75% of the rest of the days of the year; a 2% chance that there will be rain on 50% of the rest of the days of the year; a 1% chance that there will be rain on 25% of the rest of the days of the year; and there is a 0% chance that there will be no rain for the rest of the year.

   A. what is the Bayesian Prediction probability of rain on day 11.

   (5 marks)

   B. what is the Maximum a Posterior (MAP) probability of rain on day 11.

   (5 marks)

X	Y	Class
T	T	+
T	F	-
T	F	+
T	T	+
F	T	-

Table 2: X and Y Classification Data

3. (a) In the context of machine learning, 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) Discuss the advantages and disadvantages of k**-Nearest Neighbour** classification.

(10 marks)

(d) Table 2 provides a classification for a data set of X Y pairs.

(i) Calculate the **entropy** for this classification.

(5 marks)

(ii) Calculate the **information gain** for X and Y.

(5 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) The following sets express the mappings between predicates r, p, q, s, $class1$ and $class2$: $r \rightarrow \{a1, a2, a5, a6\}$, $p \rightarrow \{a2, a3, a5, a7\}$, $q \rightarrow \{a1, a2, a6\}$, $s \rightarrow \{(a2, f), (a1, 1), (a6, f)\}$, $class1 \rightarrow \{a2\}$, $class2 \rightarrow \{a2, a6\}$.

      (i) Given the above sets give a specialisation of the rule $class1(X) \leftarrow r(X) \wedge p(X)$ such that the rule is only satisfied by $class1$ members.

      (5 marks)

      (ii) Given the above sets give a rule that will correctly classify only members of $class2$.

      (5 marks)

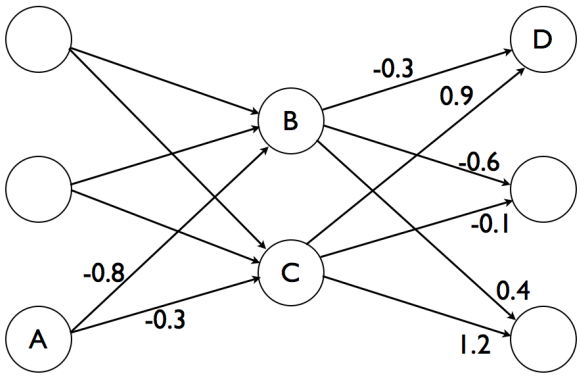

Figure 1: Example Neural Net