S228/406

DUBLIN INSTITUTE OF TECHNOLOGY
KEVIN STREET, DUBLIN 8

BSc (Hons) in Computer Science

Stage 4

SEMESTER 2 EXAMINATIONS 2011

ARTIFICIAL INTELLIGENCE II

Dr. John Kelleher
Dr. D. Lillis
Dr. I. Arana

Friday 20^{th} May
4:00 p.m. to 6:00 p.m.

Duration: 2 Hours

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

any 2 Other Questions (30 marks each).

1. (a) Explain what is meant by **inductive learning**.

(5 marks)

(b) In the context of machine learning, distinguish between **supervised** and **unsupervised** learning.

(5 marks)

(c) Inductive machine learning is often referred to as an **ill-posed problem**. What is meant by this description?

(10 marks)

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

(20 marks)

Table 1: Example feature vectors for animal classification. A 1 indicates the animal possesses the feature listed in the column, and 0 indicates they do not. The rightmost column lists the classification of each ainmal.

Species	Births Live Young	Lays Eggs	Feeds Offspring Own Milk	Warm-Blooded	Cold-Blooded	Land and Water Based	Has Hair	Has Feathers	Class
Cat	1	0	1	1	0	0	1	0	Mammal
Frog	0	1	0	0	1	1	0	0	Amphibian
Squirrel	1	0	1	1	0	0	1	0	Mammal
Duck	0	1	0	1	0	1	0	1	Bird

Table 2: The attributes of a newly discovered animal. A 1 indicates the animal possesses the feature listed in the column, and 0 indicates they do not. The column on the right contains a ? because the animal has not yet been classified.

Species	Births Live Young	Lays Eggs	Feeds Offspring Own Milk	Warm-Blooded	Cold-Blooded	Land and Water Based	Has Hair	Has Feathers	Class
Mystery	0	1	0	0	0	1	0	0	?

2. (a) You are working as an assistant-biologist to the Charles Darwin on the Beagle voyage. You are at the Galápagos Islands and you have just discovered a new animal that has not yet been classified. Table 2 lists the attributes of the animal you have found. Mr. Darwin has asked you to classify the animal using a nearest-neighbour approach and he has supplied you with a case-base of already classified animals, see Table 1.

   (i) A good measure of distance between two instances with categorical features is the number of features which have different values (the **overlap metric**, also known as the **hamming distance**). Using this measure of distance compute the distances between the mystery animal and each of the animals in the case base.

(5 marks)

(ii) If you used *1-NN* classification what class would be assigned to the mystery animal?

(5 marks)

(iii) If the you used *4-NN* classification what class would be assigned to the mystery animal?

(5 marks)

(b) Table 3 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)

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

Table 3: X and Y Classification Data

(c) The following sets express the mappings between predicates r, p, q, s, $class1$ and $class2$:

- $r \rightarrow \{a1, a2, a5, a6\}$,
- $p \rightarrow \{a2, a3, a5, a7\}$,
- $q \rightarrow \{a1, a2, a6\}$,
- $s \rightarrow \{(a2, f), (a1, 1), (a6, f)\}$,
- $class1 \rightarrow \{a2\}$,
- $class2 \rightarrow \{a2, a6\}$.

Given these 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)

Table 4: 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

3. (a) Given the joint distribution for X and Y listed in Table 4 calculate:

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

(5 marks)

(b) In you 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, and 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)

(c) Consider the following time keeping patterns of the lecturers in your college:

- 25% of lecturers start 75% of their lectures on time and 25% late.
- 50% of lecturers start 50% of their lectures on time and 50% late.
- 25% of lecturers start 25% of their lectures on time and 75% late.

(i) Given that both the 1^{st} and 2^{nd} Artificial Intelligence lectures of the year started on time, compute the posterior probability that your Artificial Intelligence lecturer follows each of the three time keeping patterns.

(10 marks)

(ii) Given that both the 1^{st} and 2^{nd} Artificial Intelligence lectures of the year started on time, what is the Bayesian Prediction that the 3^{rd} Artificial Intelligence lecture will start on time?

(5 marks)

4. (a) The following model is commonly used for continuous prediction tasks:

$$y(x) = w_0 + w_1 x_1 + \ldots + w_D x_D$$

(i) Provide the name for this model and explain all terms.

(5 marks)

(ii) Briefly describe a technique for finding optimal values for the terms $w_0, w_1, \ldots, w_D$ in the model based on a historical training set.

(5 marks)

(b) 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 $[0.1, 0.9, 0.1]$. 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., $f(x) = \frac{1}{1+\exp^{-x}}$):

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

(5 marks)

(ii) Calculate the error for each output unit.

(5 marks)

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

(10 marks)

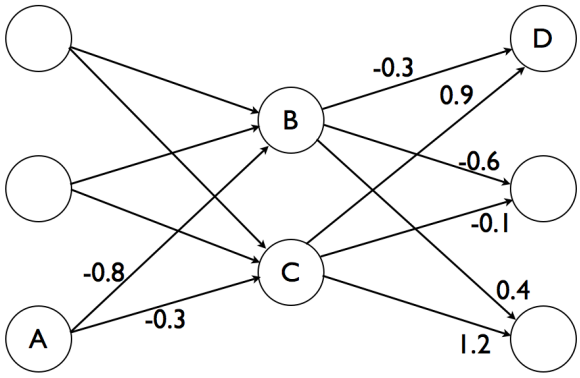

Figure 1: Example Neural Net