

R228/406

DUBLIN INSTITUTE OF TECHNOLOGY  
KEVIN STREET, DUBLIN 8

---

# **BSc (Hons) in Computer Science**

**Stage 4**

---

**SUPPLEMENTAL EXAMINATIONS 2011**

---

**ARTIFICIAL INTELLIGENCE II**

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

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, explain what is meant by **overfitting** the training data.  
(5 marks)
- (c) In the context of machine learning, explain what is meant by the term **inductive bias** and illustrate your explanation using examples of inductive biases used by machine learning algorithms.  
(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 animal.

Species	Birhs 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	Birhs 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  $1\text{-}NN$  classification what class would be assigned to the mystery animal.

(5 marks)

- (iii) If the you used  $4\text{-}NN$  classification what class would be assigned to the mystery animal.

(5 marks)

- (b) In the context of Decision Tree Learning define what is meant by the following terms:

- (i) entropy

(5 marks)

- (ii) information gain

(5 marks)

- (c) The FOIL inductive logic programming algorithm is constructing a new rule with head  $p(Y) \leftarrow$ . Which of the following literals could be considered as candidate extensions  $q(Y)$ ,  $r(X)$ ,  $s(X, Y)$ ,  $\neg s(X, Y)$ ?

(5 marks)

3. (a) Given that  $P(a|b) = 0.5$ ,  $P(a) = 0.3$ ,  $P(b) = 0.4$  calculate  $P(b|a)$ .

(5 marks)

- (b) 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 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) 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.

(10 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.

What is the Maximum a Posterior (MAP) probability of rain on day 11?

(5 marks)

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

$$y(x) = w_0 + w_1x_1 + \dots + w_Dx_D$$

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

(5 marks)

- (ii) Explain how the following model can overcome some of the limitations of the model given above.

$$y(x) = \sum_{j=0}^{M-1} w_j \phi_j(x)$$

(5 marks)

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

(5 marks)

- (c) Describe the processing stages of a McCulloch-Pitts "unit".

(7 marks)

- (d) Figure 1 is a schematic of a 3 input perceptron. Input  $a_0$  is fixed at  $a_0 = -1$ , inputs  $a_1$  and  $a_2$  are binary. The perceptron uses a threshold activation function that outputs a 1 if the weighted sum of inputs is greater than 0 and a 0 otherwise. Define the **truth-table of the function** that this perceptron implements *and* identify the **name of the function**.

(8 marks)

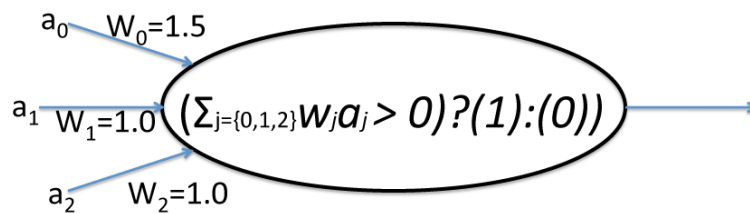


Figure 1: A 3 input perceptron. Input  $a_0 = -1$ , inputs  $a_1$  and  $a_2$  are binary. The perceptron uses a threshold activation function that outputs a 1 if the weighted sum of inputs is greater than 0 and a 0 otherwise.