







EXAM QUESTION PAPER

| | | | |
|--|---|---------------|----------------------|
| College/ Institute | Engineering, Design and Physical Sciences | | |
| Department | Computer Science | | |
| Exam Author(s) | Allan Tucker | | |
| Module Code | CS3002 | | |
| Module Title | Artificial Intelligence | | |
| Month | April/May | Year | 2020 |
| Exam Type | Full | Format | WISEflow: FLOWassign |
| Duration | <p>3 Hours plus 5 minutes to allow for upload of your work and hand in.</p> <p>Please ensure you click the green hand in button to submit your work as shown below, on the right</p> <div> <div>  1. Paper </div> <div>  2. Cover sheet </div> <div>  3. Hand in </div> </div> <div> <div>  </div> <div>  <small>Preview the cover sheet</small> </div> <div>  <small>Submission deadline: +273 days</small> </div> </div> | | |
| Number of questions | 10 Questions | | |
| Question Instructions | Answer ALL Questions All questions carry equal marks | | |
| Can students include drawings/ diagrams? | Yes | | |
| Any permitted reference materials (including external websites) | This is an open book examination | | |
| Contact for Academic Queries: | Please use the following Collaborate session link: https://eu.bbcollab.com/guest/2cc68fb26a2744fe8a4bfbebd897373f If you cannot access the Collaborate session, please email: Allan.Tucker@brunel.ac.uk or computerscience@brunel.ac.uk | | |
| Contact for technical issues: | Please email WISEflowhelp@brunel.ac.uk or use the chat tool that is embedded in WISEflow. | | |

By continuing beyond this point, you confirm that you have read the information and instructions above, and understand the conditions of this examination.

1 Unsupervised learning

- a) Calculate Euclidean distance between the following two datapoints:

$$X = [2.4, 1.0, 0.5, 3.2]$$

$$Y = [1.4, 0.0, 2.0, 1.7]$$

[3 marks]

- b) Describe how this distance can be used in K-Means Clustering to identify clusters and how it differs from Hierarchical Clustering.

[5 marks]

- c) Name one advantage of K-Means Clustering and advantage of Hierarchical Clustering.

[2 marks]

2 Supervised Learning

- a) Given the following five datapoints measuring characteristics of electric bikes using two variables and a class (representing performance of bikes):

X1: (weight = 1, top_speed = 4, class = HIGH);

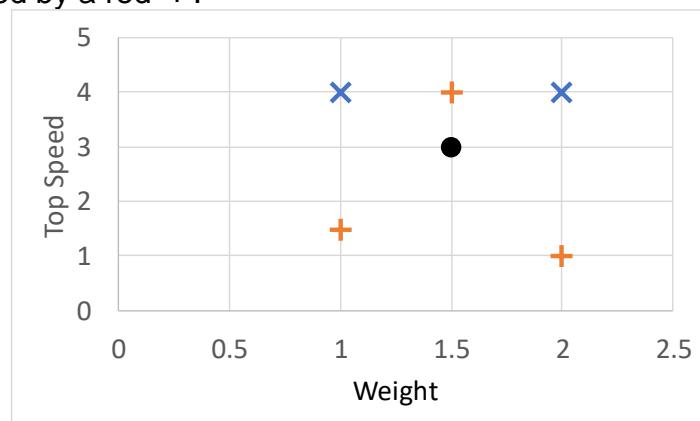
X2: (weight = 1.5, top_speed = 4, class = LOW);

X3: (weight = 1, top_speed = 1.5, class = LOW);

X4: (weight = 2, top_speed = 1, class = LOW);

X5: (weight = 2, top_speed = 4, class = HIGH);

The data can be visualised in the following graph where the HIGH performance class is represented by a blue 'x' and the LOW class is represented by a red '+'.
The new data point X6 is marked as a black circle.



Please show, *with a clear explanation*, what the classification of the following new data point would be (marked as a black circle on the chart):

X6:(weight = 1.5, top_speed = 3)

when using K Nearest Neighbour (KNN) with a simple majority vote and Euclidean distance if:

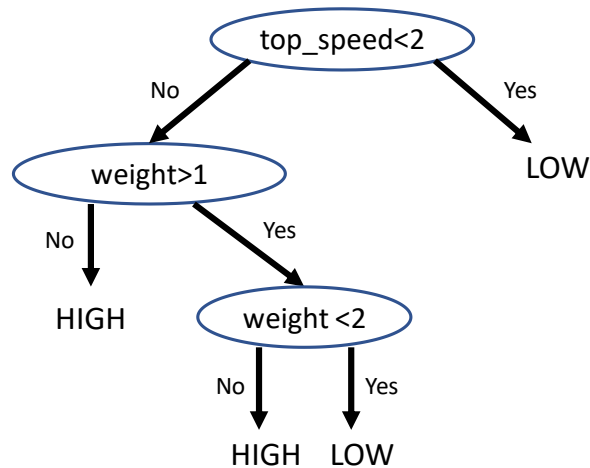
K=1

[2 marks]

K=3

[2 marks]

b) Given the Decision Tree (DT) below for classifying the same electric bike data:



How would the following new pieces of data be classified:

i) X6(weight = 1.5, top_speed = 5)

[2 marks]

ii) X7(weight = 2, top_speed = 1.5)

[2 marks]

c) Describe one advantage and one disadvantage of KNN and DTs.

[2 marks]

3 Sensitivity Analysis

a) Given the following confusion matrix:

| | | Observed | |
|-----------|-------|----------|------|
| | | False | True |
| Predicted | False | 7 | 4 |
| | True | 3 | 6 |

i) Calc Sensitivity.

[2 marks]

ii) Calc Specificity.

[2 marks]

iii) Calc Accuracy.

[2 marks]

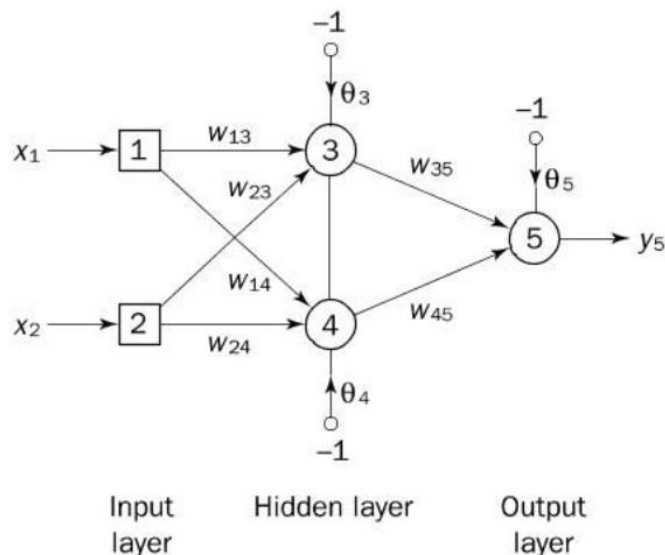
- b) Briefly describe bootstrapping and cross validation and explain why they are both useful when building or assessing machine learning classifiers. [4 marks]

4 Expert Systems

- a) Briefly describe how an expert system models human expertise and how it differs from machine learning approaches. [4 marks]
- b) Name one advantage and one disadvantage of Expert Systems over Machine Learning methods. [2 marks]
- c) What is the difference between Forward and Backward Chaining? [2 marks]
- d) Briefly describe what conflict resolution is in an expert system. [2 marks]

5 Neural Nets

- a) Given the Neural Network below with associated parameters (and no thresholds):



$W_{13} = 0.0$, $W_{23} = 0.5$, $W_{14} = -0.5$, $W_{24} = 0.5$, $W_{35} = -0.1$, $W_{45} = 1.0$, $\theta_3 = 0.2$, $\theta_4 = 0.1$, $\theta_5 = 0.2$

Calculate the output for following inputs:

- i. $x_1 = 1$, $x_2 = 0$

[3 marks]

ii. $X_1 = 1$, $X_2 = 1$

[3 marks]

- b) Explain the difference between a multilayer neural network and a perceptron and describe an example problem where the neural network will succeed but the perceptron will fail?

[4 marks]

6 Deep Learning

- a) Briefly describe what is meant by deep learning, and explain one advantage and one disadvantage associated with it.

[4 marks]

- b) Give one example of an application to which convolutional neural networks are best suited, and one example to which recurrent neural networks are best suited.

[2 marks]

- c) Explain what is meant by dropout and what is its role in a convolutional neural network.

[4 marks]

7 Deep Learning & NLP

- a) Briefly describe what Natural Language Processing (NLP) is and give one example of an NLP task.

[2 marks]

- b) Explain what one-hot encoding is and how it is used to train the word2vec embedding.

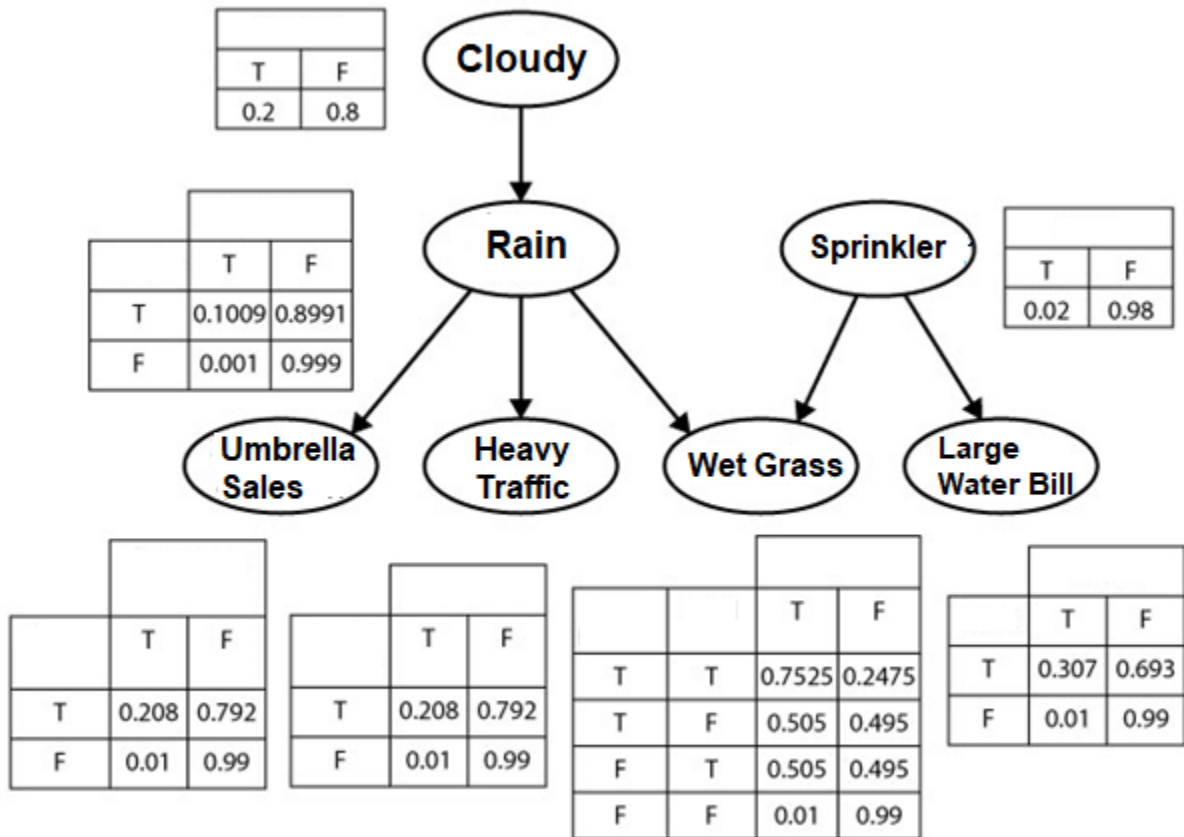
[4 marks]

- c) What is TF-IDF? Why do we include the IDF factor, instead of just TF?

[4 marks]

8 Bayesian Networks

Given the following Bayesian Network:



- a) Observing which nodes as evidence would make **Rain** conditionally independent of **Large Water Bill** [2 marks]
- b) Calculate to six decimal places, the probability of:
- Cloudy=T, Rain=F, UmbrellaSales=F, HeavyTraffic = F, WetGrass=T, Sprinkler=T, LargeWaterBill=T [3 marks]
 - Cloudy=F, Rain=F, UmbrellaSales=F, HeavyTraffic = F, WetGrass=F, Sprinkler=F, LargeWaterBill=F [3 marks]
- c) Describe one advantage of Bayesian Networks over Neural Networks. [2 marks]

9 Hidden Markov Models

a) Given the following Markov Chain:

Initial state distribution: $\pi = \begin{matrix} x & y & z \\ [0.8 & 0.2 & 0.1] \end{matrix}$

Transition Distribution: $A = \begin{matrix} & \begin{matrix} x & y & z \end{matrix} \\ \begin{matrix} x \\ y \\ z \end{matrix} & \begin{bmatrix} 0.8 & 0.2 & 0.0 \\ 0.0 & 0.9 & 0.1 \\ 0.0 & 0.0 & 1.0 \end{bmatrix} \end{matrix}$

What is the probability of the following sequences:

- i) X X X
- ii) Y Z Y
- iii) X Y Z

[6 marks]

b) Briefly describe how a hidden Markov Model differs from a Markov Chain

[2 marks]

c) Describe two examples of where a Hidden Markov Model can be used to model data

[2 marks]

10 Ethics & Society

a) Describe the Trolley Problem and explain what its implications are for AI in Society.

[5 marks]

b) Describe the Turing Test and why Searle's Chinese Room is an argument against the test proving "Strong AI"

[5 marks]