King's College London

This paper is part of an examination of the College counting towards the award of a degree. Examinations are governed by the College Regulations under the authority of the Academic Board.

Examination Period	May 2022 (Period 2)
Module Code	6CCS3ML1 and 6CCS3PRE
Module Title	Machine Learning and Pattern Recognition

Format of Examination	Written questions
Start time	30 May 10.00am
Time Allowed	1.5 hours
Instructions	You are permitted to access any materials you wish, but this is not mandated and is not expected. You may use a calculator if you find this helpful.
Rubric	ANSWER ALL QUESTIONS

The rubric for this paper must be followed and extra answers should not be submitted. For answers that are handwritten, write with blue/black ink on light coloured paper. Include the Module code, question number and student number on every page to be submitted. For answers that are typed, use the template provided.

Submission Deadline	**11.30am**
Submission Process	Work must be submitted to the **level 6** Informatics Assessments KEATS page.

Your work must be submitted as a PDF file. If you have prepared some answers on computer, and some on paper (which have then been digitised), you may upload at most two PDF files – one for computer-prepared answers, one for digitised answers. Do not duplicate answers across the two PDFs – if you do this, the computer-prepared answer will be taken. You should check that your work displays correctly after it has been uploaded.

ACADEMIC HONESTY AND INTEGRITY

Students at King's are part of an academic community that values trust, fairness and respect and actively encourages students to act with honesty and integrity. It is a College policy that students take responsibility for their work and comply with the university's standards and requirements. Online proctoring / invigilation will not be used for our online assessments. By submitting their answers students will be confirming that the work submitted is completely their own. Misconduct regulations remain in place during this period and students can familiarise themselves with the procedures on the College website

Important: Students should copy out the following statement and include it with their submission for each examination:

I agree to abide by the expectations as to my conduct, as described in the academic honesty and integrity statement.

1.

a. Describe what we mean by training error and generalisation error of a machine learning model.

[6 marks]

b. You are given a machine learning method that can search very large hypothesis spaces, and another that can search comparatively smaller hypothesis spaces. Discuss trade-offs in training and generalisation error for each of the two methods. Also consider the effect of noisy training sets.

[12 marks]

c. You have a supervised machine learning model that takes inputs $\mathbf{x}$ and computes its output $h_{\mathbf{w}}(\mathbf{x})$ using a weight vector $\mathbf{w}$. Using a training set $(\,(\mathbf{x}_1, y_1)\,(\mathbf{x}_2, y_2)\ldots(\mathbf{x}_m, y_m)\,)$, the model aims to minimise the following loss function:

$$E(h_{\mathbf{w}}) = \sum_{i=1}^{m} (y_i - h_{\mathbf{w}}(\mathbf{x}_i))^2 \tag{1}$$

i. What kind of problems can we solve with this model?

[3 marks]

ii. Give three real-world example tasks for which you can use this model.

[3 marks]

iii. How would you define your hypothesis function $h_{\mathbf{w}}(\mathbf{x})$ if your model is linear and multivariate?

[4 marks]

iv. Modify the loss function E such that you explicitly penalise complex hypotheses h. Justify your answer and explain all parameters.

[10 marks]

QUESTION 1 CONTINUES ON NEXT PAGE

SEE NEXT PAGE

v. We are now told that it is important to train the model such that the learned parameters $\mathbf{w}$ have $\|\mathbf{w}\| \simeq 1$. How might you modify the same loss function E to achieve this? Justify your answer and explain all parameters.

[12 marks]

2.

a. You are given the 3 scenarios below. For each of these, determine the type of learning you would use to solve them – supervised learning, unsupervised learning, reinforcement learning – and *briefly* explain your answer. Also write down the name of an appropriate machine learning algorithm you would choose for each scenario.

i) A news aggregator receives articles from various online news outlets. It wants to automatically analyse the news articles and to group similar articles.

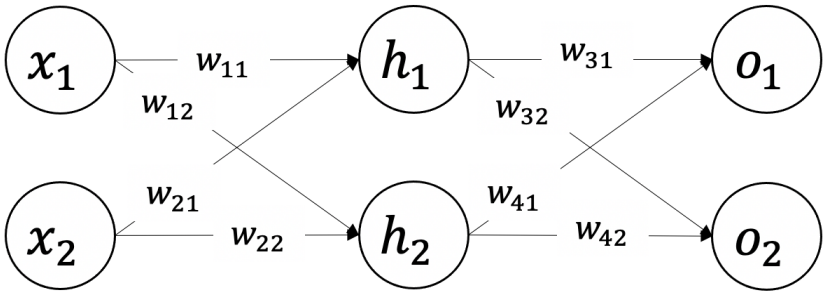
ii) An artist has an album of hand drawn images depicting different types of houses. Each image is annotated as being either a "house", "cabin", "barn", "igloo", or "villa". The artist wants to use this collection of images to build a system, which, given a new hand drawn image, determines the type of house shown in the image.

iii) A computer science student wants to develop an algorithm that can play the game "Snake". In the game, the player controls a creature that resembles a snake. The snake moves around in the game area, picks up apples, tries to avoid hitting its own tail or the walls of the game area. Each time the snake eats an apple, it gets 10 points and its tail grows longer. The player can control the direction of the snake's head with the arrow keys (up, right, down, left). The final score is based on the number of apples eaten by the snake.

[15 marks]

QUESTION 2 CONTINUES ON NEXT PAGE

b. The following diagram represents a feed-forward neural network with one hidden layer:



The input nodes to this network are nodes x_1 and x_2, and the output nodes are nodes o_1 and o_2. The weights in the network are as follows:

$w_{11} = 3$	$w_{12} = -4$	$w_{21} = -2$	$w_{22} = 5$
$w_{31} = -1$	$w_{32} = 1$	$w_{41} = 1$	$w_{42} = -1$

Each hidden and output node uses the following activation function:

$$g(s) = \begin{cases} 1, & s > 0 \\ -1, & s < 0 \end{cases}$$

Calculate the network output for the following input. Do not include biases. Show your calculations.

x_1	x_2
0	1

[4 marks]

c. You are building an artificial neural network and need to decide on the activation function for the hidden layers. You are planning to use back-propagation to train the network. What properties are required by the activation function?

[4 marks]

d. What might happen if you use the sigmoid function in a neural network with many hidden layers? Briefly explain your answer.

[2 marks]

3.

a. Consider a genetic algorithm, where you have the following two parents:

1	1	0	1	0	0	1

0	1	1	0	1	1	0

Given the following *crossover mask*:

0	1	0	1	1	0	1

Write down the resulting offspring.

[7 marks]

b. Consider a population with 4 candidates:

Candidate 1 has a fitness of 2
Candidate 2 has a fitness of 3
Candidate 3 has a fitness of 6
Candidate 4 has a fitness of 9

Use probabilistic selection to calculate the probability of selecting each candidate. Show your calculations.

[5 marks]

c. Consider a real-valued genetic algorithm where an individual has length 4 and is represented as $x = abcd$. Each of a, b, c, and d can be any digit (i.e., 0, 1, . . . , 9). Let the fitness function f be defined as

$$f(x) = (a + b) - (c + d)$$

and let the initial population consist of the following two candidates:

$$x_1 = 5703$$
$$x_2 = 1487$$

Compute the fitness of each candidate. Show your calculations.

[6 marks]

QUESTION 3 CONTINUES ON NEXT PAGE

d. Assuming the 2 individuals defined above, x_1 and x_2, represent the entire population, explain the importance of the mutation operator. *Hint*: Consider the maximum value that can be obtained using the given fitness function.

[7 marks]