

# EEEEB4023/ECEB463 Artificial Intelligence and Neural-fuzzy Systems

Dr Mohd Zafri Bin Baharuddin, Dr.

Midterm Test, Sem 2 2021/2022

1 March 2022

100 marks in this Test contribute to 20% of the final grade. Instructions:

1. This is an **OPEN BOOK** test.
2. Prepare your own pieces of paper, A4 or foolscap.
3. Redraw any required diagrams on your own paper.
4. **You have 90 MINUTES.**
5. **There are SIX (6) QUESTIONS. Answer all questions.**
6. Convert your answers to PDF using Adobe Scan, CamScanner or similar app or scanner.
7. Submit your PDF file to the Teams assignment page.

---

## **QUESTION 0 [DECLARATION]**

- (a) Rewrite the following on the first page of your submission:

*“I promise that I have not given or received aid in this test, and that I have done my part in ensuring that others as well as I uphold the UNITEN Student Rules and Regulations.”*

- (b) Below the declaration, write your full name, student ID and signature.

## **QUESTION 1 [20 MARKS]**

For each of the following, provide an answer with justifications.

- (a) Explain the use of training dataset and test dataset in machine learning. **[5 marks]**
- (b) Does the  $k$ -nearest neighbour algorithm require more computation during training time or test time? **[5 marks]**
- (c) Explain the pros and cons of the linear classifier for image classification algorithms. **[5 marks]**
- (d) In a neural network, is “activation function” a hyperparameter? **[5 marks]**

## QUESTION 2 [30 MARKS]

A training dataset consisting of four 2x2 images with its class labels is shown in the table below. A new test image is to be evaluated using the L1 Manhattan distance metric and listed in column  $d_i$ . The L1 distance equation is provided as:

$$d_i(I_{test}, I_{train,i}) = \sum_p |I_{test}^p - I_{train,i}^p|$$

where  $i$  is the data number in the training dataset, and  $p$  is the pixels of each image.

	$i$	class	image		$d_i$
Test image	–	–	30	10	–
			33	70	
Training images	0	train	84	91	$d_0$
			99	74	
	1	rocket	220	189	$d_1$
			203	157	
	2	rocket	217	188	541
			146	133	
	3	train	33	56	96
			28	28	

- Determine the distance values for  $d_0$  AND  $d_1$ . [10 marks]
- Using  $k$ -nearest neighbour, what is the output class when  $k = 3$ ? [4 marks]
- What is the output class when  $k = 2$ ? [3 marks]
- What is the output class when  $k = 1$ ? [3 marks]
- Using the Euclidean distance metric, determine the distance value for  $d_0$ . The Euclidean distance equation is:

$$d_i(I_{test}, I_{train,i}) = \sqrt{\sum_p (I_{test}^p - I_{train,i}^p)^2}$$

[10 marks]

### **QUESTION 3 [10 MARKS]**

A linear classifier has the function  $f(x, W) = Wx + b$ , with pre-trained weights and biases values as below:

$$W = \begin{bmatrix} 0.5 & 0.8 & -0.1 & 0.9 \\ 0.8 & -0.3 & 0.7 & 0.2 \\ 0.2 & -0.7 & 0.9 & 0.2 \\ 0.6 & 0.2 & -0.2 & 0.7 \end{bmatrix} \quad b = \begin{bmatrix} 0.3 \\ 0.4 \\ 0.6 \\ 0.5 \end{bmatrix} \quad classes = \begin{bmatrix} phenomenal \\ awesome \\ successful \\ generous \end{bmatrix}$$

Using the last four digits of your Student ID number as the input column vector,  $x$ , determine the output scores of the linear classifier above. Lastly, state the class that was categorised by the model.

**[10 marks]**

### **QUESTION 4 [20 MARKS]**

A *butterfly* image is input to a 3-class classification model. The raw scores of the classifier, and the correct 1-hot encoding scores are given below.

$$Score = \begin{bmatrix} \text{bee} \\ \text{beetle} \\ \text{butterfly} \end{bmatrix} = \begin{bmatrix} -3.4 \\ 1.0 \\ 7.7 \end{bmatrix} \quad Correct\_Score = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

- (a) Determine the SVM loss. **[8 marks]**
- (b) Determine the cross-entropy loss. **[8 marks]**
- (c) Is this a good or bad classifier? Why? **[2 marks]**
- (d) If all the scores were doubled, such that:

$$Score = \begin{bmatrix} \text{bee} \\ \text{beetle} \\ \text{butterfly} \end{bmatrix} = \begin{bmatrix} -6.8 \\ 2 \\ 15.4 \end{bmatrix}$$

Explain what would happen to the SVM and cross-entropy losses? Calculations not required.

**[2 marks]**

### **QUESTION 5 [10 MARKS]**

Answer the following questions based on your experience in Group Project 1.

- (a) List down the classes of your dataset. **[1 marks]**
- (b) Briefly explain the application your group submitted for Group Project 1. **[2 marks]**
- (c) In your opinion, which single hyperparameter was the most important to be modified before training? Why is that so? **[3 marks]**
- (d) A model was trained on the Teachable Machine, and its confusion matrix and accuracy per class is shown below. Comment on this model's performance based on these output results. **[4 marks]**

Class 1	39	0	0	0	0	0
Class 2	0	31	0	2	0	0
Class 3	0	0	33	0	1	0
Class 4	0	0	0	34	2	0
Class 5	0	0	0	4	31	3
Class 6	0	0	0	1	4	33
	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6

**Prediction**

CLASS	ACCURACY	# SAMPLES
Class 1	1.00	39
Class 2	0.94	33
Class 3	0.97	34
Class 4	0.94	36
Class 5	0.82	38
Class 6	0.87	38

### **QUESTION 6 [10 MARKS]**

Weight values,  $W$ , in a neural network model need to be optimized by minimizing the loss at every iteration. When a small change,  $h$ , is introduced to a weight value, the Loss rate of change can be calculated using the numerical gradient function below.

$$\frac{df(W)}{dW} = \lim_{h \rightarrow 0} \frac{f(W + h) - f(W)}{h}$$

where loss,  $L = f(W)$ . With the initial loss  $f(W) = 1.7918$  and learning rate,  $h = 0.001$ , some gradients for a single step were calculated as below.

Weight number, $W_i$ $i$	Loss with small change, $h$ $(f(W+h))$	Gradient $dL/dW$
0	1.7925	0.74
1	1.7953	3.54
2	1.802	10.24
3	1.7846	$dL_3/dW$
4	1.7914	-0.36
5	1.795	3.24

- (a) Determine the missing gradient,  $dL_3/dW$  in the table above. [6 marks]
- (b) Among the gradients in the table above, which weight number,  $i$ , should be changed to quickly reduce the loss? Please explain your chosen answer. [4 marks]

**-END OF QUESTION PAPER-**