香 港 中 文 大 學

The Chinese University of Hong Kong

二0二一至二二年度上學期科目考試

Course Examination 1 st Term, 2021-22

科目編號及名稱
Course Code & Title : CSCI3230 Fundamentals of Artificial Intelligence

時間 小時 分鐘
Time allowed : 2 hours : 0 minutes

學號 座號
Student I.D. No. : ..... Seat No. : .....

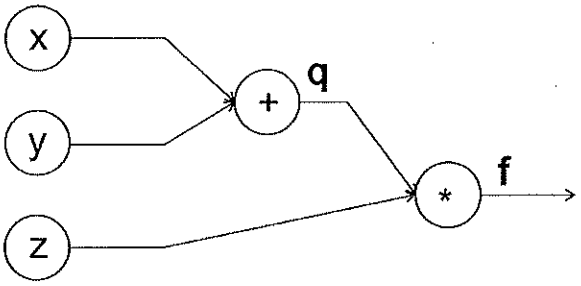
Answer **ALL** Questions. Full Score is 100%.

- This is a **close-book** examination. You can only bring one A4 page of notes as reference.
- You are **not allowed** to use electronic devices to perform calculations, and you are **not allowed** to communicate with anyone directly or indirectly during the examination.

Part I: Answer the following multiple choice questions (36%)

Part II: Answer the short answer questions (32%)

1. Consider the simple neural network as follows, suppose $x = -2, y = 5, z = -4$, what are the gradients $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial z}$ using back-propagation? (4%)



2. A single-state search problem consists of which components? List all the components and explain each of them. (4%)

3. What is the full name of PCA? What is PCA used for? Write down the full algorithm of PCA. (8%)

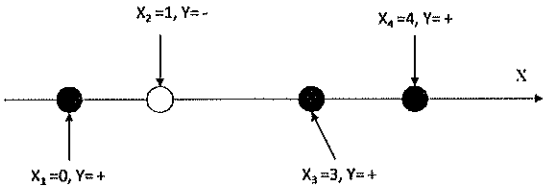
4. Why convolutional neural networks can obtain higher accuracy than traditional methods for image recognition? Does deeper models always give better performance? Why? (6%)

5. State the difference between supervised learning and unsupervised learning? Give two example methods/algorithms for each of them. (6%)

6. State two application fields of AI and give a concrete example for each field. (4%)

Part III: Answer the following analysis questions (32%)

1. (16%) A robot has collected data with its sensor. We want to use the data to build a classifier using support vector machine. Currently, the feature space is a one dimensional space $X \in R$. The desired classification output is $Y = \{+, -\}$, as shown in the figure below. The training set contains three positive examples, $x_1 = 0, x_3 = 3, x_4 = 4$, and one negative example $x_2 = 1$.



a) Currently, the data points are not linearly separable. We want to define a transformation that maps the data into a projected space in R^2. If we consider the feature mapping function as $\Phi(X) = (X, (X-1)^2)$, draw the data points after the transformation to the 2D space, and draw the separation plane. (6%)

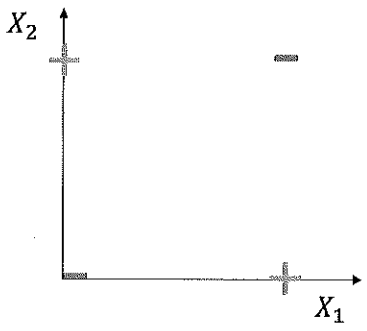
b) Indicate which examples out of x_1, x_2, x_3, x_4 are support vectors? (2%)

c) If the robot got one more negative data point $x_5 = 1.5$, would it affect the margin? Please justify the reason. (4%)

d) If the robot continues to collect a large amount of data from its sensor, SVM classifier may no longer be effective to perform the classification task, due to its limited capacity to model complex functions. In this case, you will consider to build a neural network to solve this task. Then, to design the neural network model, what aspects do you need to consider? (State at least two aspects and explain why that is important?) (4%)

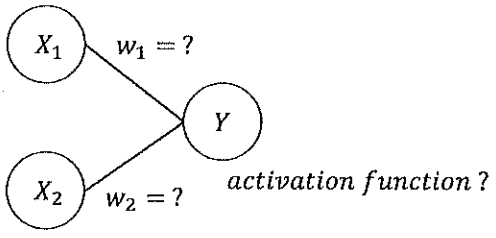
2. (16%) Recall that the XOR function for two variables has the following truth table and figure. This function is apparently non-linear. Let's try to build a neural network to mimic the function.

X_1	X_2	Y
0	0	0
0	1	1
1	0	1
1	1	0

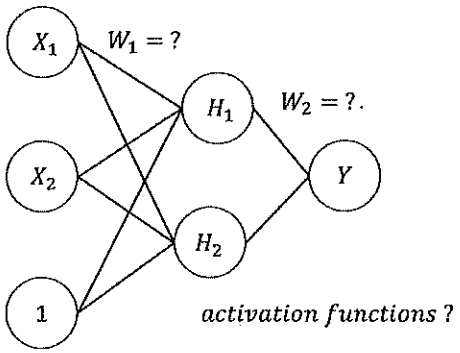


a) Can we use the following two-layer (i.e., one input layer and one output layer) network to solve it? If yes, give the weights and activation functions in the network. If no, explain why.

(6%)



b) Can we use the following three-layer network (i.e., one input layer, one hidden layer and one output layer) to solve it? If yes, give the weights and activation functions in the network, and justify why it can represent the XOR function. If no, explain why.
(Note that W_1 , W_2 denote weight matrix).
(10%)



<END OF PAPER>