University Number: _____

THE UNIVERSITY OF HONG KONG

FACULTY OF ENGINEERING
DEPARTMENT OF COMPUTER SCIENCE

DASC7606 Deep Learning
(Subclasses A, B & C)

Date: Monday, May 8, 2023 **Time: 6:30 p.m. – 8:30 p.m.**

Answer ALL questions. They are all COMPULSORY.

The mark value of each question (or part of a question) is indicated before the question (or part of the question).

Please write your answers on this examination paper in the space provided.

Only approved calculators as announced by the Examinations Secretary can be used in this examination. It is candidates' responsibility to ensure that their calculator operates satisfactorily, and candidates must record the name and type of the calculator used on the front page of the examination script.

Brand and Type of Calculator: _____

(33 pts) Question 1: Training and Data Preparation

(a) (9 pts) You are asked to design a deep learning system to perform certain tasks. For each task, determine which evaluation metric, Precision or Recall, is appropriate. Explain your choice.

(i) Detect driver fatigue to prevent any car accidents.
Answer:

(ii) Screen thousands of job applications for interview.
Answer:

(iii) Detect and raise an alarm for fire.
Answer:

(b) (9 pts) You are training a neural network for a classification task. After training for a number of epochs, the following is observed. Explain what you would do in each situation.

(i) Training converges, but the training error is high/very high.
Answer:

(ii) Training converges and training error is low, but testing error is high/very high.
Answer:

(iii) Training error does not converge but increases.
Answer:

(c) (6 pts) There are a few situations where we use backpropagation to compute the gradient of some function with respect to the pixels of an image and modify the pixels instead of the weights of the network. Describe in words (equations are not needed) what function is used for two of these situations.

Answer: (i)

(ii)

(d) (9 pts) Consider the Q-learning Algorithm. The update Q function is given below:

$$y_t = r_t + \max_a Q(s_{t+1}, a_t)$$
$$Q(s_{t+1}, a_{t+1}) \leftarrow Q(s_t, a_t) - \alpha(y_t - Q(s_t, a_t))$$

List 3 mistakes in the above equations.

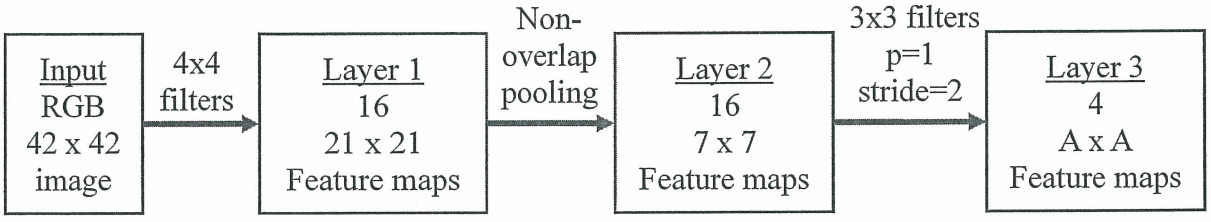
Answer:

(i)

(ii)

(iii)

(38 pts) Question 2: Convolutional Networks



The above network architecture is for sliding window object detection, each element in the AxA map represents a sliding window with 4 possible answers (no object, car, motorcycle, pedestrian)

(a) (6 pts) The Input is a 42x42 RGB image. Describe how to achieve at the feature maps of Layer 1 (i.e., number of filters, padding and stride) and the number of parameters.

Answer:

(b) (4 pts) What is the computation complexity (i.e., number of multiplication operations) for computing all the Layer 1 feature maps from the Input?

Answer:

(c) (4 pts) Describe the non-overlapping pooling (i.e., filter size and stride) between Layer 1 and Layer 2.

Answer:

Question 2: Convolutional Networks (continued)

(d) (4 pts) With 3x3 filter, padding=1 and stride=2, what is A, i.e., the size of the feature map at Layer 3? What is the number of sliding windows?

Answer:

(e) (8 pts) What is the size of each sliding window at the image?

Answer:

(f) (12 pts) Please fill in the blanks, numbered from [1] to [12], of the code according to the above architecture.

```
class SlidingWindow(nn.Module):
  def __init__(self):
    super(SlidingWindow, self).__init__()
    self.conv = nn.Sequential(
      nn.Conv2d(___[1]___, __[2]___, ___[3]__, stride=_[4]__, padding=_[5]__),
      nn.MaxPool2d(__[6]___, __[7]___),
      nn.Conv2d(___[8]___, ___[9]____, __[10]____, __[11]____, ___[12]_____)
    )
```

Ans: [1]

[2]

[3]

[4]

[5]

[6]

[7]

[8]

[9]

[10]

[11]

[12]

(29 pts) Question 3: Word embeddings

Consider the following Text1:

New technology has transformed the way music is played on the phone. Eating fresh fruit is an enjoyable and healthy way to start the day. So I like to eat a fresh fruit for my breakfast every morning while enjoying playing music on my phone.

For efficiency and better performance, all the non-key words { has, the, is, on, an, and, to, so, I, a, for, my, every, while, other } are ignored, and every word and its derivatives are treated the same, i.e., "eating" and "eats" are treated same as "eat".

Thus, the Modified Text1 is,

new technology transform way music play phone <break> eat fresh fruit enjoy healthy way start day <break> like eat fresh fruit breakfast morning enjoy play music phone

(a) (6 pts) Construct the co-occurrence matrix with **window size = 2** of the following words, {eat, enjoy, fresh, fruit, music, phone, play}, based on Modified Text1.

	eat	enjoy	fresh	fruit	music	phone	play
eat							
enjoy							
fresh							
fruit							
music							
phone							
play							

(b) (5 pts) From the co-occurrence matrix, what are the possible encoding vectors for { eat, fresh, fruit, music, phone }?

Answer:

eat =

fresh =

fruit =

music =

phone =

Question 3: Word embeddings (continued)

Given Text2:

John prefers eating fresh apple to other fruits while listening to music on his Apple phone.

Assume that the embeddings of new words are context-sensitive and derived from the embedding vectors of its neighbouring words with **window size=1**, by taking the average of the embeddings of its neighbouring words.

(c) (6 pts) What are the embeddings of "apple" and "Apple", which appear in Text2?
Hint: Modified Text2: *John prefer eat fresh apple fruit listen music Apple phone*

Answer:
apple :

Apple :

(d) (6 pts) How are the words, "eat", "play" and "enjoy", related to "apple" and "Apple"? Comment on your results.

Answer:

	eat	play	enjoy
apple			
Apple			

Question 3: Word embeddings (continued)

(e) (6 pts) Given the following word vectors, X = (0, 1, 0, 0, 2, 1, 1), Y = (1, 1, 0, 1, 2, 0, 0) and Z = (2, 1, 1, 2, 0, 0, 0). How would you assign the word vectors, X, Y and Z to "orange", "song" and "happy"? Gives reasons to justify your answer.

Answer:

orange :

song :

happy :

END OF PAPER