UNIVERSITY OF LONDON

GOLDSMITHS COLLEGE

Department of Computing

B. Sc. Examination 2020

IS53024B
Artificial Intelligence

Duration: 1 hour 30 minutes

Date and time:

This paper is in two parts: part A and part B. You should answer ALL questions from part A and TWO questions from part B. Part A carries 40 marks, and each question from part B carries 30 marks. The marks for each part of a question are indicated at the end of the part in [.] *brackets.*

There are 100 marks available on this paper.

Electronic calculators must not be programmed prior to the examination. Calculators which display graphics, text or algebraic equations are not allowed.

THIS PAPER MUST NOT BE REMOVED
FROM THE EXAMINATION ROOM

Part A

Question 1

(a) i. What is meant by *underfitting*? [5]

ii. What is meant by *overfitting*? [5]

iii. Sketch a training and validation loss plot and indicate regions of under- and overfitting. [5]

iv. What is meant by *generalisation*? [5]

Question 2

Consider the following Keras deep learning model:

```
model = models.Sequential()
model.add(layers.Dense(64, activation = 'relu', input_shape = (13,)))
model.add(layers.Dense(64, activation = 'relu'))
model.add(layers.Dense(1))
```

(a) What type of problem is appropriate for the model? Justify your answer. [5]

(b) Suggest an appropriate loss function. [5]

(c) What is the meaning of `model = models.Sequential()`? [5]

(d) Explain the meaning of `activation = 'relu'`. [5]

Part B

Question 3

(a) Illustrate, with diagrams, the action of a 5 x 5 convolution filter with a stride length of one on an input image of size 28 x 28 pixels. (You need not draw every pixel, but include enough of them so that the convolutional operation is evident.) [8]

(b) How many trainable parameters are there in a 5 x 5 filter? Explain your answer. [4]

(c) Consider the following convolutional layer:

```
model = models.Sequential()
model.add(layers.Conv2D(32, (5, 5), activation = 'relu',
                        input_shape = (28, 28, 1)))
```

i. How many filters are there in the layer? [2]

ii. What is the window size of each filter? [2]

iii. How many trainable parameters are there in the layer? [4]

(d) Now consider the following convolutional model:

```
model = models.Sequential()
# layer 1
model.add(layers.Conv2D(32, (3, 3), activation = 'relu',
                        input_shape = (28, 28, 1)))
# layer 2
model.add(layers.Conv2D(64, (3, 3), activation = 'relu'))
# layer 3
model.add(layers.Conv2D(64, (3, 3), activation = 'relu'))
```

i. The 3 x 3 window in the third layer receives information from what size window in the first layer? [2]

ii. Why might this be a problem? [4]

iii. How might you solve the problem? [4]

Question 4

(a) Consider this Keras deep learning model:

```
from tensorflow.keras import models
from tensorflow.keras import layers

model = models.Sequential()
model.add(layers.Dense(32, activation = 'relu', input_shape = (10000,)))
model.add(layers.Dense(16, activation = 'relu'))
model.add(layers.Dense(1, activation = 'sigmoid'))
```

i. What is the shape of the weight tensor of the second (middle) layer? Explain your answer. [5]

ii. What transformation is performed by the second layer? [5]

(b) Explain, in the context of model training, the *mini-batch stochastic gradient descent* algorithm. Your answer should explain why the algorithm reduces model loss and the meaning of the terms 'mini-batch' and 'stochastic', but should not explain momentum. [10]

(c) i. Gradient descent is an imperfect optimisation algorithm. Why? [2]

ii. Explain, by referring to pseudocode, how *momentum* helps avoid the problem alluded to in the previous question part. [8]

Question 5

(a) *The few computers that existed in America in 1955 were mostly just used for numerical calculations — but that was about to change. As Simon would later tell it, "Over Christmas, Newell and I invented a thinking machine."*

Explain the significance of Simon and Newell's 'thinking machine': what problem it solved, and the cognitive assumption that lay behind Simon's claim. [5]

(b) What is the *combinatorial explosion*? [5]

(c) What was the early remedy to this combinatorial explosion?

(d) What is the fundamental limitation of classical AI? [5]

(e) What is the modern solution to the above limitation? Illustrate your answer with an example. [5]

(f) What is the relationship between classical AI, machine learning, and the workings of the brain? [10]