

UNIVERSITY OF LONDON

GOLDSMITHS COLLEGE

Department of Computing

B. Sc. Examination 2019

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) Copy and complete the following table of definitions. The first definition has been provided as a guide.

[20]

Term	Definition
Scalar regression	A task where the target is a single continuous value
AI	
Loss function	
Backpropagation	
Hypothesis space	
Sample and target	
Multi-label classification	
Dense layer	
Supervised learning	
Hold-out validation	
Sigmoid neuron	

Question 2

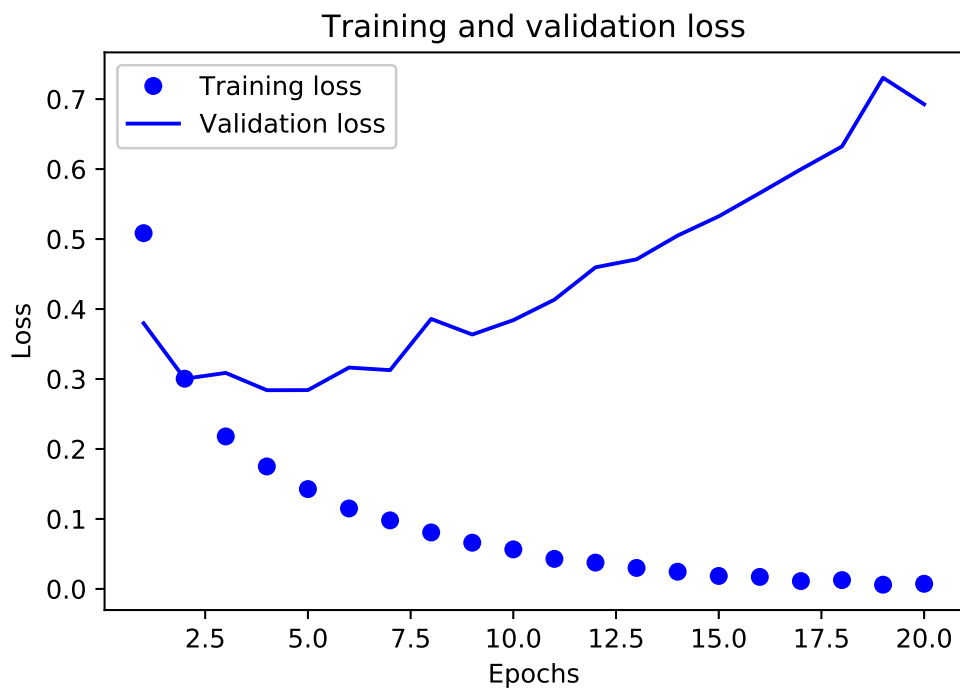
(a) Consider the following Keras deep learning model:

```
from keras import models
from keras import layers

model = models.Sequential()
model.add(layers.Dense(64, activation = 'relu', input_shape = (10000,)))
model.add(layers.Dense(64, activation = 'relu'))
model.add(layers.Dense(46, activation = 'softmax'))
```

- i. What is the intended type of classification task? Explain your answer. [4]
- ii. How many layers are there? [1]
- iii. How many output classes are there? [1]
- iv. How many trainable parameters are there? Show your working. [4]

(b) This question part concerns the figure 'Training and validation loss' .



- i. Describe and explain the shape of the training loss plot. [5]
- ii. Describe and explain the shape of the validation loss plot. [3]
- iii. What do the plots imply for the model? [2]

Part B

Question 3

- (a) Explain the convolution operation. [6]
- (b) Suppose, in an image classification task, the training image set is quite small - just 2000 images - and no further images are available. What measures could be taken to mitigate overfitting? [2]
- (c) Convolutional networks frequently employ *max pooling* layers. What is a max pooling layer and what is its purpose? [6]
- (d) The employment of a pre-trained network is a common approach to deep learning on small image datasets. Describe how a pre-trained convolutional network could be used on a small dataset. Provide a motivation for each step that you describe. [6]
- (e) Consider the following convolutional base:

```
from keras import layers
from keras import models

model = models.Sequential()
model.add(layers.Conv2D(32, (3, 3), activation='relu', input_shape=(28, 28, 1)))
model.add(layers.MaxPooling2D((2, 2)))
model.add(layers.Conv2D(64, (3, 3), activation='relu'))
model.add(layers.MaxPooling2D((2, 2)))
model.add(layers.Conv2D(64, (3, 3), activation='relu'))
```

Copy the summary table for this convolutional base and fill in the missing values (marked ??). [10]

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, ??, ??, 32)	320
max_pooling2d_1 (MaxPooling2D)	(None, ??, ??, 32)	0
conv2d_2 (Conv2D)	(None, ??, ??, 64)	18496
max_pooling2d_2 (MaxPooling2D)	(None, ??, ??, 64)	0
conv2d_3 (Conv2D)	(None, ??, ??, 64)	36928

Question 4

- (a) One hot encoding is frequently used in text processing. Suppose we have a Python dictionary:

```
{1: 'slasher',
 2: 'worst',
 3: 'truly,',
 4: 'exaggerating,',
 5: 'of',
 6: 'ever',
 7: 'this',
 8: 'one',
 9: 'is',
10: 'made',
11: 'movies',
12: 'without',
13: 'the'}
```

What is a one-hot encoding of the phrase: 'worst ever made'? Show your working. [4]

- (b) What is text embedding? [4]
- (c) Why are feedforward networks generally not recommended for sequential data? [2]
- (d) Recurrent neural networks (RNNs), on the other hand, are a common choice for the prediction of sequence data. Explain how the simple recurrent unit, as implemented in a Keras layer by `SimpleRNN`, functions. Illustrate your answer with pseudocode. [6]
- (e) An RNN implemented with a Keras layer of `SimpleRNN` units does not perform well in practice. Why is that? [4]
- (f) Explain how the long short-term layer helps address the above problem. Illustrate your answer with pseudocode and ensure that each element of the pseudocode is explained. [10]

Question 5

- (a) Write a short account on the application of deep learning to retro Atari console games. [10]
- (b) Write an account of the four key technologies that have contributed to the deep learning revolution. [10]
- (c) Discuss the limitations of deep learning. [10]