UNIVERSITY OF BIRMINGHAM

School of Computer Science

Third Year Undergraduate

20416

Neural Computation

Main Summer Examinations 2019

Time allowed: 1:30

[Answer all questions]

Note

Answer ALL questions. Each question will be marked out of 20. The paper will be marked out of 60, which will be rescaled to a mark out of 100.

Question 1

(a) Categorise each of the four machine learning tasks below as either unsupervised learning, supervised learning, or reinforcement learning. Note that some machine learning tasks can be solved by combining different approaches, e.g. semi-supervised learning. Here, you should select only ONE category for each machine learning task.

1. You are given a set of movie reviews, where each review consists of a text in French describing a movie, as well as a rating between 1 and 6 "stars" of that movie. Your task is to predict the movie rating, given only the review text of a new movie.
2. Without any training data, you are asked to partion a set of pictures of clothes into different groups, such that the clothes within a group are similar to eachother, while clothes from different groups are dissimilar.
3. You have access to a large number of emails, where each email has been tagged as either "spam" or "no spam". Your task is to learn to distinguish spam emails from no spam emails given the email content.
4. Learn how to play the PacMan computer game.

Copy the table below in your answer sheet, and mark your answers with tick marks.

	Unsupervised L.	Supervised L.	Reinforcement L.
1			
2			
3			
4			

[4 marks]

(b) State one definition of machine learning. State one example of a machine learning application which matches this definition. Justify your answer.

[6 marks]

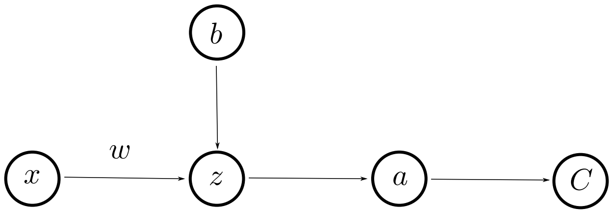

Figure 1: A machine learning model with two parameters w and b.

(c) Consider the simple machine learning model in Figure 1 which has two parameters $w, b \in \mathbb{R}$, one input unit x, and one output unit $a = \sigma(z)$, where $z = wx + b$, and $\sigma : \mathbb{R} \to \mathbb{R}$ is some differentiable, non-linear function. We are given m elements of training data $(x_1, y_1), \ldots, (x_m, y_m)$, where for each element $i \in \{1, \ldots, m\}$, x_i is an input, and y_i is the corresponding output.

We plan to train the model by minimising the cost function

$$\tilde{C}(w, b) = \left(\sum_{i=1}^{m} C^{(i)}(w, b) \right) + w^2$$

where for each $i \in \{1, \ldots, m\}$,

$$C^{(i)}(w, b) := \frac{1}{2} \sum_{i=1}^{m} (y^{(i)} - \sigma(wx^{(i)} + b))^2.$$

Assume we initialise the model with parameters $w = w_0 \in \mathbb{R}$ and $b = b_0 \in \mathbb{R}$. What are the values of the model parameters w and b after one iteration of gradient descent with learning rate $\varepsilon > 0$?

[10 marks]

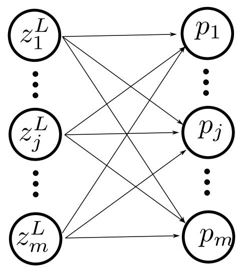

Figure 2: A softmax layer.

Question 2

(a) State the pseudo-code for the gradient descent algorithm with learning rate ε. Explain how this algorithm is used in neural computation.

[4 marks]

(b) For what type of machine learning task can a softmax layer be useful? In a softmax layer with m units, the output unit p_j can be defined as

$$p_j := \frac{e^{z_j^L}}{\sum_{k=1}^{m} e^{z_k^L}},$$

where z_j^L is the input from previous layers, as shown in Figure 2. However, a naive implementation of softmax can lead to numerical problems, i.e., "overflow". Explain why this problem can occur, and how to remedy it. **[6 marks]**

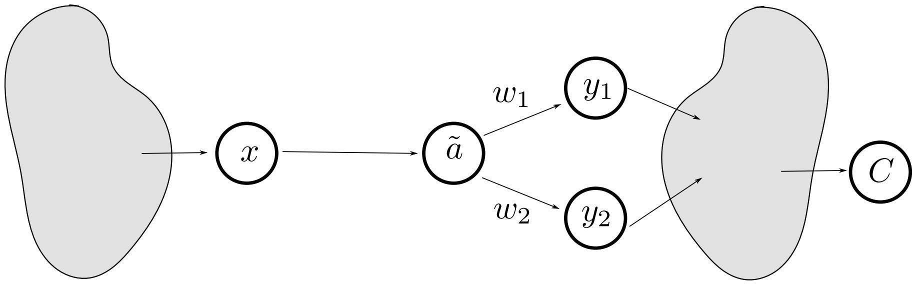

Figure 3: Part of a neural network with a dropout activation unit $\tilde{a}$.

(c) What is "dropout", and why is it used in neural computation? Figure 3 shows part of a neural network where C is a cost function, and $y_1 := w_1\tilde{a}$ and $y_2 := w_2\tilde{a}$. Assume that no other units than y_1 and y_2 take input from unit $\tilde{a}$. We want the unit $\tilde{a}$ to be a dropout activation unit from the unit x. How would you define $\tilde{a}$? Assume that you already know $\frac{\partial C}{\partial y_1}$ and $\frac{\partial C}{\partial y_2}$, e.g., by using the back-propagation algorithm. What is $\frac{\partial C}{\partial x}$? **[10 marks]**

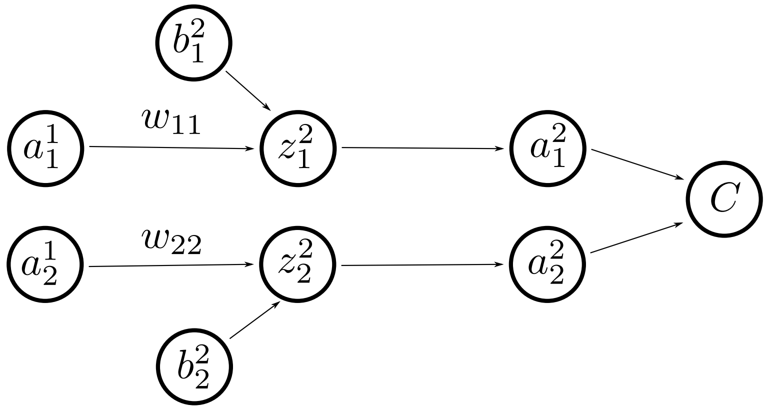

Figure 4: Neural network in Question 3 (c).

Question 3

(a) Give one example of how the structure of a biological neural network reflects the structure of the world.

[4 marks]

(b) State the universal approximation theorem. What assumptions does the theorem make? What is the significance of the theorem for neural computation? What are the limitations of the theorem?

[8 marks]

(c) Consider the feed forward neural network depicted in Figure 4 which has two input units a_1^1 and a_2^1, and two output units a_1^2 and a_2^2, and cost function C. For $j \in \{1, 2\}$, we have the following relationships between the units

$$\begin{aligned} z_j^2 &= w_{jj} a_j^1 + b_j^2 \\ a_j^2 &= \phi(z_j^2), \text{ and} \\ \frac{\partial C}{\partial a_j^2} &= (a_j^2 - y_j), \end{aligned}$$

where ϕ is some differentiable activation function which satisfies $\phi(x) \in (0, 1)$ and $\phi'(x) = \phi(x)(1 - \phi(x))$ for all $x \in \mathbb{R}$.

Compute the partial derivatives $\frac{\partial C}{\partial w_{11}}$ and $\frac{\partial C}{\partial b_1^2}$.

Now, assume that the inputs satisfy $a_1^1 > a_2^1 > 0$, and with the current weights and biases, the output unit satisfies $a_1^2 > y_1$ where (y_1, y_2) is the correct output on input (a_1^1, a_2^1). We will apply gradient descent with learning rate $\varepsilon > 0$ to update the weights and biases of the network. Will the weight parameter w_{11} increase? Justify your answer.

[8 marks]

This page intentionally left blank.

Do not complete the attendance slip, fill in the front of the answer book or turn over the question paper until you are told to do so

Important Reminders

- Coats/outwear should be placed in the designated area.
- Unauthorised materials (e.g. notes or Tippex) must be placed in the designated area.
- Check that you do not have any unauthorised materials with you (e.g. in your pockets, pencil case).
- Mobile phones and smart watches **must** be switched off and placed in the designated area or under your desk. They must not be left on your person or in your pockets.
- You are not permitted to use a mobile phone as a clock. If you have difficulty seeing a clock, please alert an Invigilator.
- You are not permitted to have writing on your hand, arm or other body part.
- Check that you do not have writing on your hand, arm or other body part – if you do, you must inform an Invigilator immediately
- Alert an Invigilator immediately if you find any unauthorised item upon you during the examination.

Any students found with non-permitted items upon their person during the examination, or who fail to comply with Examination rules may be subject to Student Conduct procedures.