

**Xi'an Jiaotong-Liverpool University**

**西交利物浦大學**

PAPER CODE	EXAMINER	DEPARTMENT	TEL
CSE315		CSSE	

**1st SEMESTER 2018/19 FINAL EXAMINATIONS**

**BACHELOR DEGREE – Year4 / MASTER DEGREE – Year 1**

**MACHINE LEARNING**

**TIME ALLOWED: 2 Hours**

---

**INSTRUCTIONS TO CANDIDATES**

- 1) Total marks available are 100. This exam account for 70% of the final mark.
- 2) Answer all questions.
- 3) The number in the column on the right indicates the marks for each question.
- 4) Answers should be written in the answer booklet(s) provided.
- 5) The university approved calculator - Casio FS82ES/83ES can be used.
- 6) All the answers must be in English.

**Question 1**

(a) Machine learning is the discipline of making a machine to learn from experience  $E$  and carry out a set of tasks  $T$  through incremental improvement of performance measure  $P$ . Give a concrete machine learning example with specific  $E$ ,  $T$  and  $P$ .

(4 Marks)

(b) Explain the difference between supervised learning and unsupervised learning, and give one example for the unsupervised learning.

(4 Marks)

(c) Algorithms/models can be designed based on either (i) traditional computer science techniques or (ii) machine learning techniques. Compare the two approaches.

(4 Marks)

**Question 2**

For predicting the housing prices in a specific region, a model can be built based on previous sales records. Read the data in Table 1, and answer the following questions.

Table 1. Sales records

<b><math>x</math>: number (bedroom)</b>	<b><math>y</math>: price (\$1,000)</b>
2	447
3	643
5	1048
4	846
6	1246
7	1444
9	1896
3	649
5	1021
3	646

(a) Based on data visualization, two hypotheses  $y_1 = \theta_0 + \theta_1 x$  and  $y_2 = \theta_0 + \theta_1 x + \theta_2 x^2$  can be proposed. Indicate which parameters should be calculated for linear regression, and present the corresponding cost function.

(4 Marks)

(b) The Gradient Descent algorithm can be applied to solve the prediction task. (i) Describe the process of updating parameters in two hypotheses, and give the formula. (ii) If each of the parameters is initialized as 0.1 and the learning rate is set as 0.01, present the parameters after one iteration using the Gradient Decent algorithm.

(9 Marks)

(c) Describe what criterion can be applied to compare the predicting performance of these two hypotheses, and explain which hypothesis will be chosen eventually.

(5 Marks)

### Question 3

(a) List two activation functions for neural networks, and present the range of output for each of them.

(4 Marks)

(b) For a given Multi-Layer Perceptron as illustrated in Figure 1, if 9 attributes are used as its input, compute the amount of parameters.

(3 Marks)

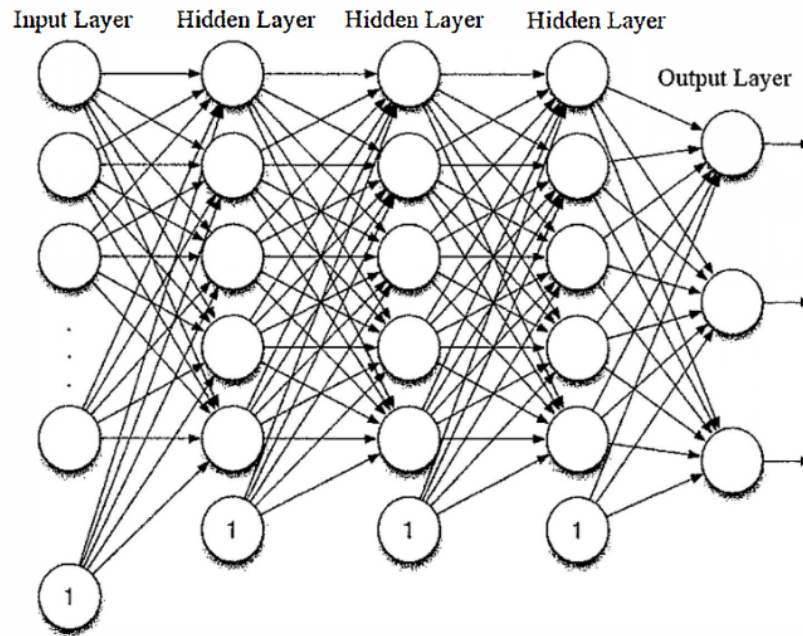


Figure 1. A Multi-Layer Perceptron

(c) Describe the processes of forward propagation and backpropagation. Given a neural network in Figure 2, if learning rate is set as 0.5, use (i) forward propagation and (ii) backpropagation to update the value of  $w_5$ .

(10 Marks)

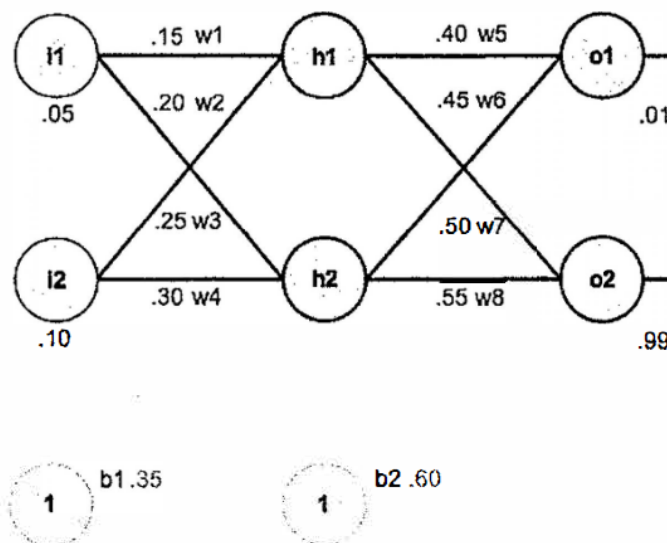


Figure 2. A Neural Network

(d) Figure 3 shows a simple Convolutional Neural Network. Explain the operations in each of the convolutional layer, pooling layer, and fully-connected layer. Give an application example that involves the convolutional neural network.

(7 Marks)

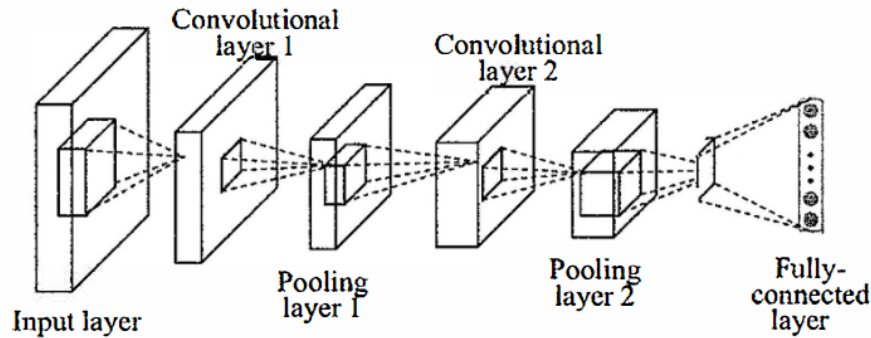


Figure 3. A Convolutional Neural Network

(e) Figure 4 shows an LSTM model, which is a kind of Recurrent Neural Network model. Explain the components (forget gate, input gate, output gate, and cell state) in an LSTM cell.

(6 Marks)

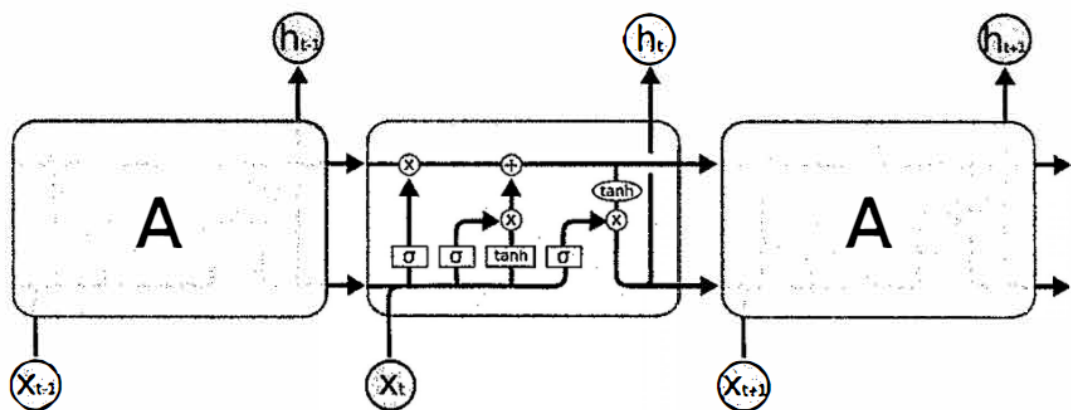


Fig. 4 An LSTM Model

**Question 4**

Clustering is a commonly-used unsupervised learning algorithm. Based on the records shown in Table 2, answer the questions below.

Table 2. Sample records

No.	$x_1$	$x_2$
1	1	1
2	1.5	2
3	3	4
4	5	7
5	3.5	5
6	4.5	5
7	3.5	4.5

(a) The records in Table 2 can be divided into 2 clusters based on  $x_1$  and  $x_2$ . The k-means algorithm and Euclidean distance are employed for this task, where Records #1 and #4 are chosen as the seeds. (i) Present the process of the two iterations of the k-means algorithm's execution, and (ii) give the clustering result and the centers of the two clusters after two iterations.

(10 Marks)

(b) Summarize the weaknesses of the k-means algorithm.

(5 Marks)

**Question 5**

What is the Principal Component Analysis (PCA)? Explain this algorithm and describe its main application.

(10 Marks)

**Question 6**

Suppose that you are working for a social media company, and your task is to analyze the user's emotion (e.g., happiness, sadness, anger, etc.) for a given post (including texts, pictures and videos) in the past.

(a) What data preparation is needed at first?

(3 Marks)

(b) What machine learning techniques can be applied in this task? List three algorithms/models, and explain how to apply them.

(12 Marks)

**END OF EXAM PAPER**

**THIS PAPER MUST NOT BE REMOVED FROM THE EXAMINATION ROOM.**