

The University of Nottingham Ningbo China

SCHOOL OF COMPUTER SCIENCE

A LEVEL 3 MODULE, AUTUMN SEMESTER 2019-2020

MACHINE LEARNING

Time allowed: 2 hours

Candidates may complete the front cover of their answer book and sign their desk card but must NOT write anything else until the start of the examination period is announced

Answer All Questions

Only silent, self contained calculators with a Single-Line Display or Dual-Line Display are permitted in this examination.

Dictionaries are not allowed with one exception. Those whose first language is not English may use a standard translation dictionary to translate between that language and English provided that neither language is the subject of this examination. Subject specific translation dictionaries are not permitted.

No electronic devices capable of storing and retrieving text, including electronic dictionaries, may be used.

DO NOT turn examination paper over until instructed to do so

Question 1:

- (a) Discuss the concept of **classification** and **regression** and provide concrete examples when each method is appropriate.

[8 marks]

- (b) Referring to the lecture notes, what is deep learning (with respect to representation learning)? Briefly describe the two problems occur when we go deeper, vanishing gradient and parameter explosion. You may answer the first question by drawing a typical deep learning architecture with explanation.

[10 marks]

Question 2:**[20 marks]**

- (c) We will use a single ADLINE classifier to classify 2D real number input into two classes (+1 and -1). The training data is listed as follows.

Input	Output
(1,2)	1
(-2,3)	1
(1,-2)	-1
(-2,-1)	-1

Train the ADLINE with online mode using the training data above from the first row to the last row, with the initialization of weights $w=[1 \ 0 \ 0]^T$ and $\eta=0.5$. For each iteration, calculate the current output of ADLINE (o) and the updated weights (w). Note that you need to show intermediate steps of your calculation. An answer without the intermediate steps will not receive any marks

[16 marks]

- (d) What is the difference between online mode and batch mode when training ADLINE?

[4 marks]

Question 3:**[24 marks]**

In a laparoscopic surgical tool detection task, it is known that the surgical tool components come from one of two classes, C_1 or C_2 . Each sample of a surgical tool component Y has two features, $Y = (y_1, y_2)$. An experiment has collected 6 samples and their feature values and classification are shown in the following table. Answer the following questions by showing details of intermediate steps. An answer without the intermediate steps will not receive any marks.

$Y = (y_1, y_2)$		Classification
y_1	y_2	
3	3	C_2
8	8	C_1
3	5	C_2
5	4	C_2

6	5	C ₁
7	6	C ₁
6	3.5	=?

- (a) Based on the data with known classification, use a KNN classifier with K=3 by the distance metric to classify the unknown sample Y=(6, 3.5) in the last row of the table.

[8 marks]

- (b) Suppose the data in the above table are collected in a way that the classification is unknown, i.e., we do not know C in the last column. Cluster the data into two clusters using K-Means algorithm with initial cluster centers Z₁=(3, 3) and Z₂=(5, 4).

[16 mark]

Question 4:

[38 marks]

- (a) Briefly describe the steps to design a learning system according to the lecture notes.

[10 marks]

- (b) Briefly describe the idea of soft margin SVM. What is the problem it tries to tackle and how the problem is solved.

[6 marks]

- (c) The formula used for Bayesian learning is as follows. Describe what do the terms (variables) in the formula correspond to in the context of machine learning.

$$P(h|x) = \frac{P(x|h)P(h)}{P(x)}$$

[10 marks]

h	Hypothesis, class label
x	
$P(h)$	
$P(x)$	
$P(x h)$	
$P(h x)$	

- (d) LSTM is an excellent model for solving sequence-to-sequence problems. It contains multiple components such as forget gate, input gate, output gate, etc. Briefly describe the functionality of forget gate and input gate.

[6 marks]

- (e) Most the state-of-the-art deep architectures consist of pre-training phase followed by fine-tuning phase. Please briefly describe what is pre-training and fine-tuning and Why it helps with the model's performance.

[6 marks]