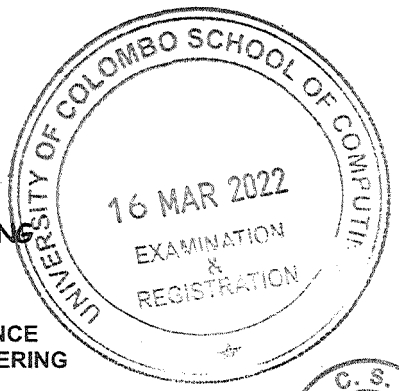
UNIVERSITY OF COLOMBO, SRI LANKA

UNIVERSITY OF COLOMBO SCHOOL OF COMPUTING



BACHELOR OF SCIENCE IN COMPUTER SCIENCE
BACHELOR OF SCIENCE HONOURS IN COMPUTER SCIENCE
BACHELOR OF SCIENCE HONOURS IN SOFTWARE ENGINEERING

Academic Year 2021/2022 – Third Year Examination – Semester I

SCS3201 – Machine Learning and Neural Computing

TWO (2) HOURS (for both Part A & B)

126

- Part A -

To be completed by the candidate

Index No: .....

Important Instructions to candidates:

1. The medium of instruction and question is **English**.
2. If a page or a part of this question paper is not printed, please inform the supervisor immediately.
3. Note that questions appear on both sides of the paper. If a page is not printed, please inform the supervisor immediately.
4. Write your index number on each and every page of the Question paper.
5. Students are required to answer both **Part A** and **Part B** in **two hours**.
6. Answer **ALL** questions. There are **03** questions in **Part A** and **01** Question in **Part B** of the paper.
7. **Part A** consists of **11 pages,** and it carries **65 marks**.
8. Any electronic device capable of storing and retrieving text including electronic dictionaries and mobile phones are **not allowed**.
9. **Non-Programmable** calculators are **allowed**.

For Examiner's use only	
Part A	
Question No	**Marks**
1	
2	
3	
Total	

Index No: .....

Part A

Question 1

a. Differentiate ***Regression*** and ***Classification*** tasks in ***Supervised Machine Learning***.

[03 Marks]

Answer:

b. Write a key difference in ***Unsupervised Machine Learning*** compared to ***Supervised Machine Learning***.

[03 Marks]

Answer:

c. A few samples of the zoo dataset and its data attributes are given below. The original dataset consists of 101 animals in seven (07) classes (i.e., Mammal, Bird, Reptile, Fish, Amphibian, Bug and Invertebrate). Answer the following questions using the available information in this dataset.

Animal Name	hair	feathers	eggs	milk	airborne	aquatic	predator	toothed	backbone	breathes	venomous	fins	legs	tail	domestic	catsize	Class type
antelope	1	0	0	1	0	0	0	1	1	1	0	0	4	1	0	1	1
frog	0	0	1	0	0	1	1	1	1	1	0	0	4	0	0	0	5
termite	0	0	1	0	0	0	0	0	0	1	0	0	6	0	0	0	6
penguin	0	1	1	0	0	1	1	0	1	1	0	0	2	1	0	1	2
crab	0	0	1	0	0	1	1	0	0	0	0	0	4	0	0	0	7
catfish	0	0	1	0	0	1	1	1	1	0	0	1	0	1	0	0	4
cheetah	1	0	0	1	0	0	1	1	1	1	0	0	4	1	0	1	1

i. If you are asked to classify the data into labeled classes what are the attributes that you **consider** and attributes that you **discard** for your classification task? Briefly explain the reasons for discarded attributes if any.

[03 Marks]

Answer:

ii. Assume a data set that contains **only** the given samples (07 samples in total). Write **irrelevant** attributes (apart from the two attributes you discarded above in the Question 1.c.i.) if there is any in such a case. Briefly explain reasons.

[03 Marks]

Answer:

iii. Assume a student is going to cluster the entire dataset with **K-Means** algorithm to investigate separation of these classes into different clusters. What is the most reasonable value for K in *K-Means* algorithm to try? Briefly explain your justification.

[04 Marks]

Answer:

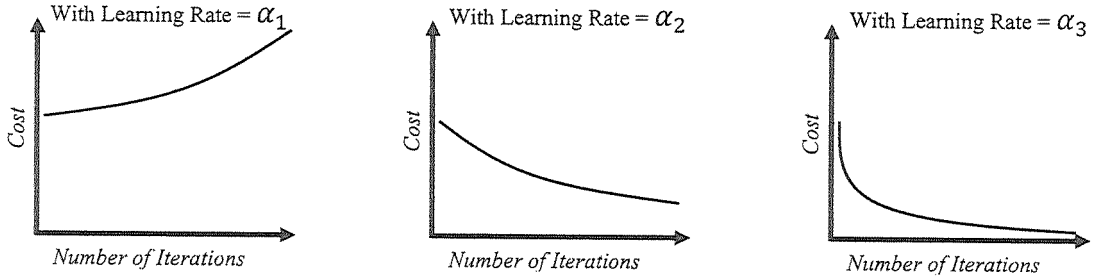
Index No: .....

d. What is the purpose of a *Cost Function* in a supervised learning task?

[03 Marks]

Answer:

e. Assume a student has run a *Linear Regression* task three times on a dataset without changing any configuration except the learning rate. The student applied three different learning rates $\alpha_1, \alpha_2, \alpha_3$ in three times and the corresponding cost function plots for three instances are given below.



Which of the three learning rates ($\alpha_1, \alpha_2, \alpha_3$) has the **highest** and **lowest** magnitude? Briefly explain your reasoning behind the judgements.

[06 Marks]

Answer:

Index No: .....

Question 2

a. Typically ***Mean Squared Error*** is not used as an evaluation metric in **Logistic Regression**. Why does **Mean Squared Error** is inappropriate as an evaluation metric in Logistic Regression?

[03 Marks]

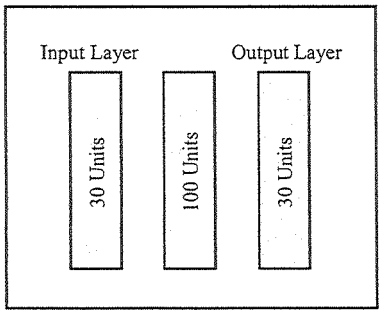
Answer:

b. It is a known fact that the *Gradient Descent algorithm* may lead to a local optimum during the learning process. With this risk of ending up in a sub-optimal solution, how appropriate to use the *Gradient Decent algorithm* during the learning process when you have a Squared Error function as the cost function? Justify your answer.

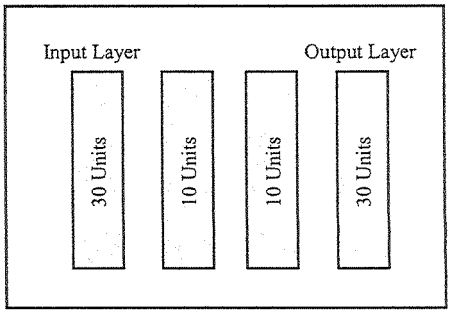
[03 Marks]

Answer:

c. Consider the following two (02) fully connected feed forward artificial neural network architectures. The architecture 1 has three (03) layers and Architecture 2 has four (04) layers.



Architecture 1



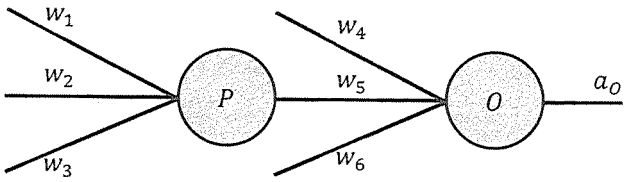
Architecture 2

Which architecture will lead to a more complex training process with respect to the **time and memory** considerations? Explain your answer with justifications.

[04 Marks]

Answer:

d. Two neurons, P and O in a multilayer feedforward artificial neural network with backpropagation learning algorithm are depicted below. Neuron O is in the output layer and the neuron P is in the layer immediately before the output layer. For a given input x with the expected value of y, the neuron O produces output a_O. The cost function is defined as E which is a squared error function.



Assuming a learning rate of α and using appropriate derivative terms, define the weight adjustment rule for weight parameter w_4 according to the gradient descent cost minimization procedure.

[02 Marks]

Answer:

e. Using appropriate derivative terms and notations, explain how you are going to derive the necessary terms in the weight adjustment rule in order **to calculate the weight adjustment** for parameter w_1 of the artificial neural network in the Question 2.d.

[05 Marks]

Answer:

f. ***k-fold cross-validation*** is often used method for machine learning model validation. It randomly partitioned the dataset into k equal sized subsets. A single subset of the k subsets is retained as the validation data for testing the model, and the remaining $k-1$ subsets are used as training data. The cross-validation process is then repeated k times, with each of the k subsets used exactly once as the validation data. The k results can then be averaged to produce a single estimation.

Write an advantage and a disadvantage of k-fold cross-validation.

[03 Marks]

Answer: ***Advantage***

Answer: ***Disadvantage***

Index No: .....

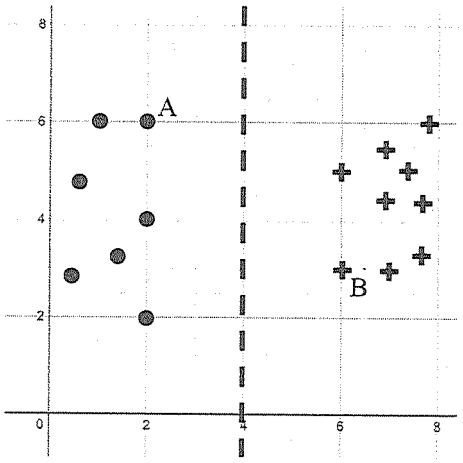
Question 3

a. Briefly explain the use of *Laplace smoothing* in Naïve Bayes classifier.

[03 Marks]

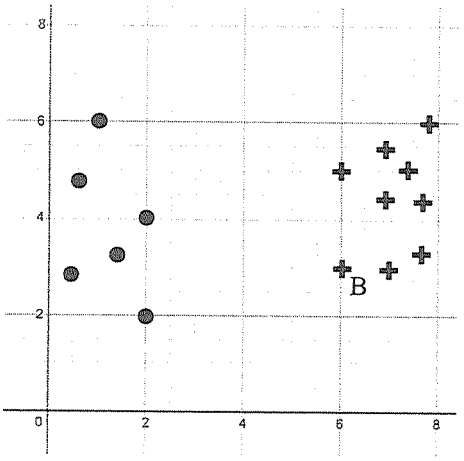
Answer:

b. Assume a **hard-margin SVM** finds its decision boundary as given in the dashed line below in the diagram.



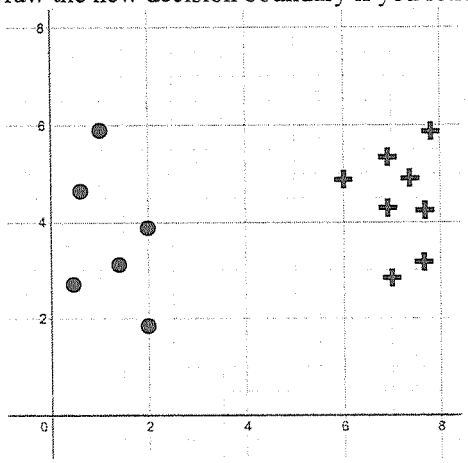
Draw the new decision boundary in the following plot, if the datapoint **A is removed** from the dataset and the SVM is retrained.

[03 Marks]



c. Let's assume the datapoint B is also removed from the dataset after removing the datapoint A in Question 3.b. Draw the new decision boundary if you retrain the SVM with the new dataset.

[03 Marks]



d. Consider the following table to answer subsequent questions in Question 3.d. The table shows a dataset of COVID-19 test results of 10 individuals against three attributes.

#	Cold	Temperature	Fatigue	Result
1	No	Normal	No	- Negative
2	No	High	Yes	- Negative
3	No	High	Yes	+ Positive
4	No	Very High	Yes	+ Positive
5	Yes	Normal	No	- Negative
6	Yes	Normal	Yes	+ Positive
7	Yes	High	No	- Negative
8	Yes	High	Yes	+ Positive
9	Yes	Very High	No	+ Positive
10	Yes	Very High	Yes	+ Positive

i. Calculate the ***Entropy*** for the dataset given above considering the target variable Result. Show your calculations clearly.

[03 Marks]

Answer:

ii. In order to decide the first split of the decision tree, it has been decided to measure the ***Entropy*** for **three possible cases**. It has been found that the ***Entropy*** for two cases as given below.

- *Entropy(Result | Cold) = 0.93*
- *Entropy(Result | Fatigue) =0.70*

Find the third case *Entropy(Result | Temperature)* by showing your calculation steps clearly.

[03 Marks]

Answer:

iii. Calculate the *Information Gain* for **all three cases** that were mentioned in the **Question 3.d.ii.** Show your calculations clearly.

[03 Marks]

Answer:

iv. Based on the above calculations in Question 3.d, what is the feature that will be used as the first split in decision tree construction. State your reasons clearly.

[02 Marks]

Answer:

****

126

UNIVERSITY OF COLOMBO, SRI LANKA

UNIVERSITY OF COLOMBO SCHOOL OF COMPUTING

UCSC

BACHELOR OF SCIENCE IN COMPUTER SCIENCE
BACHELOR OF SCIENCE HONOURS IN COMPUTER SCIENCE
BACHELOR OF SCIENCE HONOURS IN SOFTWARE ENGINEERING

Academic Year 2021/2022 – Third Year Examination – Semester I

SCS3201 – Machine Learning and Neural Computing

TWO (2) HOURS (for both Part A & B)

- Part B -

To be completed by the candidate

Index No: .....

Important Instructions to candidates:

1. The medium of instruction and question is **English**.
2. If a page or a part of this question paper is not printed, please inform the supervisor immediately.
3. Note that questions appear on both sides of the paper. If a page is not printed, please inform the supervisor immediately.
4. Write your index number on each and every page of the Question paper.
5. Students are required to answer both **Part A** and **Part B** in **two hours**.
6. Answer **ALL** questions. There are **03** questions in **Part A** and **01** Question in **Part B** of the paper.
7. **Part B** consists of **6 pages,** and it carries **35 marks**.
8. Any electronic device capable of storing and retrieving text including electronic dictionaries and mobile phones are **not allowed**.
9. **Non-Programmable** calculators are **allowed**.

For Examiner's use only	
Part B	
Question No	**Marks**
4	
Total	

Part B

Question 4

a. Suppose the two fuzzy sets F and G given below with X = {1, 2, 3, 4, 5}.

$$F = \left\{\frac{1}{1} + \frac{0.5}{2} + \frac{0.25}{3} + \frac{0.1}{4}\right\}$$

$$G = \left\{\frac{0.5}{1} + \frac{1}{2} + \frac{0.5}{3} + \frac{1}{4} + \frac{0.5}{5}\right\}$$

Answer the following.

(i) What is G' ?

[02 Marks]

Answer:

(ii) What is $F \cap G$?

[02 Marks]

Answer:

(iii) What is $F - G$?

[02 Marks]

Answer:

b. Consider the following Octave implementation of Mamdani FIS which computes the percentage of investments to be allocated to stocks using four rules.

```
[System]
Name='Investment-Portfolio'
Type='mamdani'
Version=2.0
NumInputs=2
NumOutputs=1
NumRules=4
AndMethod='min'
OrMethod='max'
ImpMethod='min'
AggMethod='max'
DefuzzMethod='centroid'

[Input1]
Name='Age'
Range=[20 100]
NumMFs=2
MF1='Young':'trapmf',[-21 20 40 80]
MF2='Old':'trapmf',[40 80 100 101]

[Input2]
Name='Risk-Tolerance'
Range=[0 10]
NumMFs=2
MF1='Low':'trapmf',[-1 0 2 8]
MF2='High':'trapmf',[2 8 10 11]

[Output1]
Name='Percentage-In-Stocks'
Range=[0 100]
NumMFs=3
MF1='About-Fifteen':'trimf',[-10 15 40]
MF2='About-Fifty':'trimf',[25 50 75]
MF3='About-Eighty-Five':'trimf',[60 85 110]

[Rules]
1 2, 3 (1) : 2
2 1, 1 (1) : 2
-2 -1, 2 (0.5) : 1
-1 -2, 2 (0.5) : 1
```

The four rules are:

1. *If (Age is Young) or (Risk-Tolerance is High), then (Percentage-In-Stocks is About-Eighty-Five) (1)*
2. *If (Age is Old) or (Risk-Tolerance is Low), then (Percentage-In-Stocks is About-Fifteen) (1)*
3. *If (Age isn't Old) and (Risk-Tolerance isn't Low), then (Percentage-In-Stocks is About-Fifty) (0.5000)*
4. *If (Age isn't Young) and (Risk-Tolerance isn't High), then (Percentage-In-Stocks is About-Fifty) (0.5000)*

Where the number inside the parenthesis represents the rule weight.

Answer the following questions.

(i) List down the **five (5) steps** in the Mamdani fuzzy inference process.

[05 Marks]

Answer:

Index No: .....

(ii) Plot the **membership functions** of the two inputs and the output with labels and legends.

[06 Marks]

Answer:

(iii) Calculate the approximate **value for *Percentage-In-Stocks*** when the two inputs *Age* and *Risk-Tolerance* are 40 and 7, respectively. Give the **aggregation graph**. State any assumptions you make.

[08 Marks]

Answer:

Index No: .....

c. What are the typical elements of a reinforcement learning system?

[03 Marks]

Answer:

d. Consider the following grid world problem.

S		**+5**

The states are grid squares, identified by their row and column number (row, col). The agent always starts in state (1,1), marked with the letter **S**. The agent receives a reward of +5 if it reaches the terminal goal state (1,3). The agent can move in four directions (North, East, South, and West).

Give the **reward table** for this problem. State any assumptions you make.

[07 Marks]

Answer:

****