Department of Computer Science and Engineering

University of Asia Pacific (UAP)

Program: B.Sc. in Computer Science and Engineering

Final Examination	**Spring 2020**	**4th Year, 1st Semester**
Course Code: CSE 427	**Course Title: Machine Learning**	**Credits: 3.00**
Full Marks: 120* (Written)		**Duration: 2 Hours**

* Total Marks of Final Examination: 150 (Written: 120 + Viva: 30)

Instructions:
1. There are **Four (4)** Questions. Answer all of them. All questions are of equal value. Partial marks are shown in the margins.
2. Non-programmable calculators are allowed.

---

1. **a)** Suppose we have 20 data points belonging to three classes. 7 data points belong to class A, *X* belong to class B and the rest belong to class C. [10]

Here, X = last digit of your ID +1

Calculate the entropy of the dataset.

b) "K -means clustering is a supervised algorithm and k-NN algorithm is an unsupervised algorithm." – do you agree with this statement? Explain why or why not. [5]

c) Suppose we have a dataset as follows— [15]

x	y
2	17
4	28
6	42

We want to apply linear regression to predict the value of y. Our hypothesis function is:
$$h(\theta) = \theta_0 + \theta_1 x$$

At initial step, initialize the values of θ_0 and θ_1 as the last two digits of your ID. For Example,
    If your ID is 14101026, then $\theta_0 = 2$ and $\theta_1 = 6$.
    If your ID is 14101007, then $\theta_0 = 0$ and $\theta_1 = 7$. etc.

Here, learning rate, alpha = 0.1

Now, what will be values of θ_0 and θ_1 after updating them only once using gradient decent? (Your task is to calculate the updated values of θ_0 and θ_1 after one iteration)

2. **a)** Depict overfitting and underfitting in a plot diagram. Mention some ways of how we could overcome these situations that hinders the proper tuning of a regression algorithms prediction. [8]

b) Plot the following activation functions with equation— [12]

      i.   Sigmoid Function
      ii.   Hyperbolic Tangent Function
      iii.   ReLU

c) Describe how to determine the value of 'k' in k-means clustering algorithm with the necessary plot illustration. [10]

3. **a)** Apply pooling and normalization (using ReLU) operations associated with Convolutional Neural Network for the following 2D array— [15]

0.77	-0.11	0.11	0.33	0.55	-0.11	0.33
-0.11	1.00	-0.11	0.33	-0.11	0.11	-0.11
0.11	-0.11	1.00	-0.33	0.11	-0.11	0.55
0.33	0.33	-0.33	0.55	-0.33	0.33	0.33
0.55	-0.11	0.11	-0.33	1.00	-0.11	0.11
-0.11	0.11	-0.11	0.33	-0.11	1.00	-0.11
0.33	-0.11	0.55	0.33	0.11	-0.11	0.77

Fig. 01: 2D array input feed for pooling and normalization in ConvNet.

Assume that, pooling window is n*n and stride is 2.
Where n = (Your ID mod 2) + 2.

b) Describe the relation between Receiver Operating Characteristics (ROC) and Area under the Curve (AUC) in a single plot drawing. [5]

c) Suppose you designed a prediction algorithm that can predict whether a person has COVID-19 or not. You examined 1000 people's data. Now, the prediction result and the actual result is indicated as follows— [10]
      i.   Number of persons has Covid-19 and predicted the same – 12.
      ii.   Number of persons has Covid-19 and predicted the opposite – 6.
      iii.   Number of persons does not have Covid-19 and predicted the same – 974.
      iv.   Number of persons does not have Covid-19 and predicted the opposite – 8.
Now, create the confusion matrix indicating True Positive, True Negative, False Positive and False Negative. Also determine the precision, recall and F_1 score.

4. **a)** Suppose we have a function f which follows the following truth table: [15]

x_1	x_2	x_3	f
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	0
1	1	0	0
1	1	1	1

Now, design a single neuron with appropriate weights so that the neuron takes inputs x_1, x_2, x_3, and outputs f.

b) Draw and indicate the mimication between a human neuron and a single neuron perceptron. [5]

c) Suppose, we have 1 lac data points and we are trying use SVM algorithm to classify two classes. Among the 1 lac data points, there are 100 data points which are the Support Vector points. [10]

Now, if we discard all other data points other than our Support Vector points, will it affect our algorithm anyhow?

Or,

4. **a)** Determine the name of the logic gate it simulates by the following Artificial Neural Network— [15]
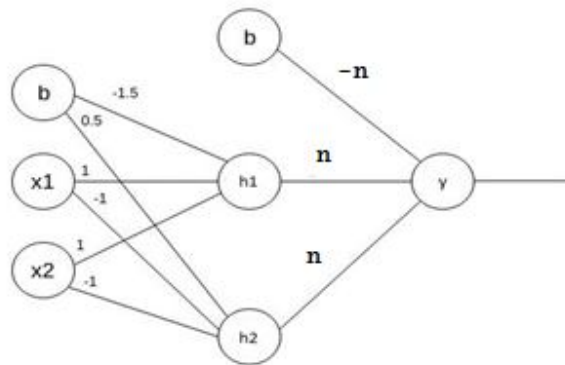

Fig. 02: Implementation of a logic gate using Artificial Neural Network.

b) "A single neuron works as the logistic regression algorithm" – do you agree with this statement? Explain why or why not. [5]

c) Determine the maximum value of the function $f(x, y, z) = x + y + 2z$, using Lagrange Multipliers method, where there is a constraint to satisfy, and the constraint is defined as $x^2 + y^2 + z^2 = A$. [10]

Where, A is one greater than the last digit of your ID.