DEDAN KIMATHI UNIVERSITY OF TECHNOLOGY

University Examinations 2020/2021

M.Sc (BUSINESS ANALYTICS) YEAR ONE SEMESTER TWO

CIT 6114: MACHINE LEARNING

DATE: **TIME:**

Instructions: Answer Question 1 and Any Other Two.

Question 1: (30 Marks)

a) Discuss the main difference between Discriminative and Generative Models **[2 Marks]**

b) By giving TWO examples of each, discuss the difference between Parametric vs. Nonparametric Machine Learning Algorithms **[5 Marks]**

c) In deep learning, the process involves both forward propagation and backpropagation, describe the two processes **[4 Marks]**

d) Using an appropriate diagram, discuss how a Neural Networks get the optimal Weights and Bias values **[4 Marks]**

e) Discuss the use of Support Vectors in support vector machines **[3 Marks]**

f) It is crucial for marketers and policy makers to be aware of different classes of clients to address their need to serve them better. Attracting new customers is no longer a good strategy for mature businesses since the cost of assisting existing customers is much lower. Discuss four types of client's segmentation **[4 Marks]**

g) Having a large number of dimensions in the feature space can mean that the volume of that space is very large, and in turn, the points that we have in that space (rows of data) often represent a small and non-representative sample. Discuss THREE methods for dimensionality reduction **[3 Marks]**

h) Using appropriate examples discuss FIVE applications of reinforcement learning **[5 Marks]**

Question 2: (15 Marks)

a) Suppose there is a marketing company A, who does various advertisement every year and get sales on that. The below list shows the advertisement made by the company in the last 5 years and the corresponding sales:

Advertisement	Sales
$90	$1000
$120	$1300
$150	$1800
$100	$1200
$130	$1380
$200	??

The company wants to do the advertisement of $200 in the year 2022 and wants to know the prediction about the sales for this year. Using regression analysis provide forecasted sales for the year **[10 Marks]**

b) Discuss FIVE different types of regression analysis used in machine learning **[10 Marks]**

Question 3: (15 Marks)

a) Consider the following data generated by classification algorithms **A** and **B** on a particular data set.

Algorithm A

Sensitivity (TP rate)	0	0.56	0.78	0.91	0.96	1
Specificity	1	0.99	0.81	0.42	0.02	0

Algorithm B

Sensitivity (TP rate)	0	0.42	0.58	0.64	0.9	1
Specificity	1	0.99	0.8	0.68	0.04	0

i. On the same grid, plot ROC curves for the two classification algorithms.

(9 marks)

ii. Interpret the graphs to identify the best performing algorithm.

(2 marks)

b) Demonstrate how **linear models** can be used in *binary classification* problems.

(4 marks)

Question 4: (15 Marks)

a) With an illustration, briefly describe the *instance-based knowledge representation* i.e., the *nearest-neighbor* classification method.

(7 marks)

b) Consider the following rules derived from contact lens data that tells the kind of contact lens to prescribe, given certain information about a patient.

Using these rules, construct a decision tree that describes the structure of the lens data.

(8 marks)

END.