National University of Computer and Emerging Sciences, Lahore Campus

Course Name:	Artificial Intelligence	Course Code:	CS 401
Program:	BS(CS)	Semester:	Fall 2017
Duration:	1 hr	Total Points:	30
Paper Date:	Friday, Oct 3 2017	Weight	20%
Section:	ALL	Page(s):	5
Exam Type:	Mid-II		

Student : Name:_____ Roll No._____ Section:_____

Instruction/Notes: Please Solve all questions on the question paper If needed you might attach rough sheets at the end.

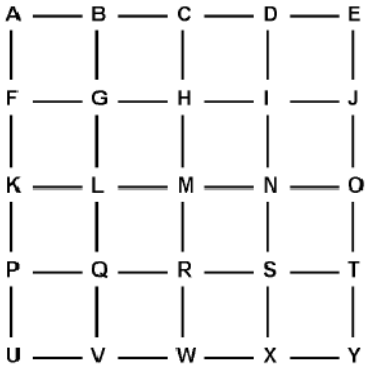
Problem 1 [Short Questions] [1.5 + 1.5 + 2 + 2 Points]

What is the main difference between classification and regression?

What is the main difference between supervised and unsupervised learning paradigm?

Name five supervised learning algorithms.

In the following diagram, A through Y are all data points. Each data point has features (x,y) corresponding to its coordinate in the grid. If data points **A** through **K** belong to class 1 and data points **L** through **Y** belong to **class 2**, what is the classification of **M**? when using the k = 5 nearest neighbors to classify it? Assume that Euclidean distance is used to compute the nearest neighbors.



Problem 2. Decision Tree [5 + 3 Points]

For this problem we will use the data set as shown below that has three binary input attributes, A, B, C, and one binary outcome/class attribute Y. The three input attributes, A, B, C take values in the set {0,1} while the Y attribute takes values in the set {True, False}.

A	B	C	Y
0	1	1	True
1	1	0	True
1	0	1	False
1	1	1	False
0	1	1	True
0	0	0	True
0	1	1	False
1	0	1	False
0	1	0	True
1	1	1	True

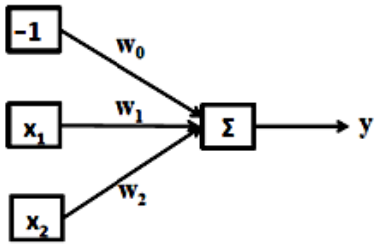
Dataset

Part a) The information Gain for features A and B have already been computed and are **0.126** and **0.093** respectively. **Your job** in this part is to **compute the information gain of C** and show a single decision-node tree that will be learned by **ID3** from the above dataset.

Part b) Compute the training accuracy of the decision tree

Problem . Perceptron Learning **[4 + 2 + 4 Points]**

Part a) Consider the perceptron with first input always set to -1 and the two features $\mathbf{X_1}$ and $\mathbf{X_2}$ are binary. The output ***y*** is also binary and is computed using a threshold as shown below



$$y = \begin{cases} 1 & \text{if } (w_0 + w_1 \cdot x_1 + w_2 \cdot x_2) > 0 \\ 0 & \text{otherwise} \end{cases}$$

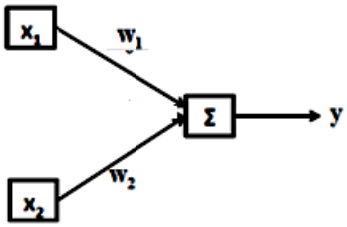
i) Which one(s) of the following choices for the weight vector [w_0 w_1 w_2] can correctly classify y as y = (X_1 AND X_2)? Here AND refers to the logical AND operation.

- o [1 1 0]
- o [-1.5 1 1]
- o [-2 1 1.5]
- o No weights can compute the logical AND relation.

ii) Which one(s) of the following choices for the weight vector [w_0 w_1 w_2] can correctly classify y as y = (X_1 XOR X_2)? Here XOR refers to the logical XOR operation.

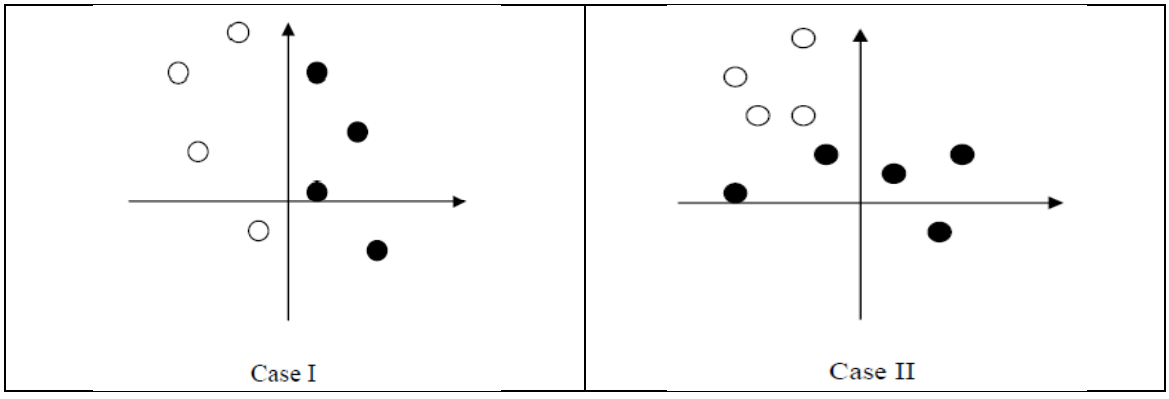
- o [1 1 0]
- o [-1.5 1 1]
- o [-2 1 1.5]
- o No weights can compute the logical XOR relation.

Part b) Now consider the a perceptron architecture without a bias term and with the two real valued features X_1 and X_2 as shown below



$$y = \begin{cases} 1 & \text{if } (w_1 \cdot x_1 + w_2 \cdot x_2) > 0 \\ 0 & \text{otherwise} \end{cases}$$

Can the network distinguish the two classes in the cases illustrated below? Why/Why not in each case?



Answer

Case 1:

Case 2:

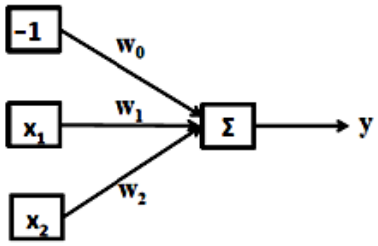
Part c) Derive a gradient descent training rule for a single unit with output o, where

$$o = w_0 + w_1 x_1 + w_1 x_1^2 + \ldots + w_n x_n + w_n x_n^2$$

Problem . Genetic Algorithms for Learning a Perceptron [3 + 2 Points]

It is well known that the weights of a perceptron can be learned using Genetic Algorithms. As in any application of GAs the first step would be the design a chromosome (representation of solution) for this problem. Once the chromosome representation is finalized we also need to define the fitness function, the crossover operator and the mutation operation.

For the perceptron with **a bias term** and two real valued features $\mathbf{X_1}$ and $\mathbf{X_2}$ as shown below



$$y = \begin{cases} 1 & \text{if } (w_0 + w_1 \cdot x_1 + w_2 \cdot x_2) > 0 \\ 0 & \text{otherwise} \end{cases}$$

Part a) Describe a detailed chromosome representation for learning it's weights using GAs.

Part b) How would you compute fitness of a your chromosome when learning the weights of this perceptron in the supervised learning setting.