Subject Name: Statistical Foundation for Machine LearningMachine Learning, End-Sem, MM:40 Marks, 4/5/24

Question 1 : [7.5 Points] Answer the following :

a) **[5 Points]:** An unbiased dice is rolled and for each number on the dice a bag is chosen:

Numbers on the Dice	Bag chosen
1	Bag A
2 or 3	Bag B
4 or 5 or 6	Bag C

Bag A contains 3 white ball and 2 black ball, bag B contains 3 white ball and 4 black ball and bag C contains 4 white ball and 5 black ball. Dice is rolled and bag is chosen, if a white ball is chosen find the probability that it is chosen from bag B.

b) [**2.5 Points**] If A and {B,C} are conditionally independent given D (e.g p(A,B,C|D) = p(A|D)p(B,C|D)), are A and B conditionally independent given D? Hint: you can use the fact $P(X) = \sum_{Y} P(X, Y)$.

Question 2 : [5 Marks] Consider two variables(X, Y) which takes two values 0 and 1, we want to sample from their joint distribution p(X, Y) which is defined as follows: p(X=0, Y=0) = 0.2, p(X=1, Y=0) = 0.3, p(X=0, Y=1) = 0.4, p(X=1, Y=1) = 0.1. Generate 5 pairs of samples of X and Y while initializing X=0 and Y=0. NOTE: Consider the sample with highest probability every iteration. HINT : Gibbs Sampling.

Question 3 [5 Marks] Consider a single sigmoid threshold unit with three inputs, x1, x2, and x3.

$y = g(w0 + w1x1 + w2x2 + w3x3)$ where $g(z) = 1 / 1 + \exp(-z)$

We input values of either 0 or 1 for each of these inputs. Assign values to weights w0, w1, w2 and w3 so that the output of the sigmoid unit is greater than 0.5 if an only if (x1 AND x2) OR x3.

Question 4 : [7.5 Points]

We have a training set consisting of samples and their labels. All samples come from one of two classes, 0 and 1. Samples are two dimensional vectors. The input data is the form {X1,X2,Y} where X1 and X2 are the two values for the input vector and Y is the label for this sample.

After learning the parameters of a Naıve Bayes classifier we arrived at the following table: Table 1: Naıve Bayes conditional probabilities

	Y=0	Y=1
X1	P(X1=1|Y =0) = 1/5	P(X1=1|Y =1) = 3/8
X2	P(X2=1|Y =0) = 1/3	P(X2=1|Y =1) = 3/4

Denote by w1 the probability of class 1 (that is w1 = P(Y = 1)). If we know that the likelihood of the following two samples: {1,1,1},{0,0,0} given our Naıve Bayes model is 1/160, what is the value of w1? Derive an explicit value for w1. Simplify as best as you can.

Question 5 : [15 Marks]

Given 15 points in the cartesian coordinate system as A1:(4,9), A2 : (9,1), A3: (3,5), A4 : (1,2), A5 : (12,5), A6 : (2,3), A7 : (14,7), A8 : (5,10), A9 : (10,12), A11 : (8,8), A12 : (4,6), A13 : (3,13), A14 : (1,8), A15 : (6,11). Kindly perform K-means clustering while considering K=3.