Q1. Consider the following dataset:

ID	X1	X2	Y	ID	X1	X2	Y
1	1	-1	1	7	6	1	1
2	2	0	1	8	1	-1	0
3	1	0	1	9	7	-3	1
4	-1	-2	1	10	-5	1	0
5	3	2	1	11	2	-2	1
6	-3	3	0	12	-3	1	0

In case of the above dataset, which of the following logistic regression models will produce the least value of the logistic regression loss function? [Logistic Regression loss function: -Y*log(σ(W.X)) - (1-Y)*log(1-σ(W.X))] where σ denotes sigmoid function.

- P(Y=1|X1,X2) = σ(X1)
- P(Y=1|X1,X2) = σ(X2)
- P(Y=1|X1,X2) = σ(X1 + X2 - 3)
- P(Y=1|X1,X2) = σ(X1 - X2 + 1)

Q2. Convert the features (X1,X2) into binary attributes (X1′,X2′) as follows: X1′=1 if X1>=0, else X1′=0, similarly for X2. Now build a Naïve Bayes classifier for the above dataset. What is the prediction and risk for the test case (-1,2)?

- 0, 2/11

Q3. In a system, each data-point has two attributes: X1 and X2. Both attributes take only integer values from 0 to 9 (both included). The label Y can take 2 values. Prior distribution: p(Y=1)=2/3, p(Y=2)=1/3.
Class conditional distributions for the attribute are defined as follows:
p(X1=x | Y=1) = 1/10 (uniform)
p(X2=x | Y=1) = k*x where k is a constant.

p(X1=x | Y=2) = k1 (if 0<=x<=4), = 2*k1 (if 5<=x<=9),
p(X2=x | Y=2) = k2 (if 0<=x<=2), = 3*k2 (if 3<=x<=9)

In the same setting as above, calculate the risk of prediction at the point (X1=2, X2=9) using a Naive Bayes Classifier

Q4. Now consider another system where X1 and X2 both can take values (0,1,2), i.e. 9 points are possible. Y can take 3 class labels: A,B,C. Apriori, all class labels are equally likely.

$p(X1=x1, X2=x2 \mid Y=A) = c1*\max(x1,x2)$ where c1 is a constant
$p(X1=x1, X2=x2 \mid Y=B) = c2*\min(x1,x2)$ where c2 is a constant
$p(X1=x1, X2=x2 \mid Y=C) = c3*|x1-x2|$ where c3 is a constant

At which feature vectors will a Bayesian classifier (not Naive) predict Y to be A? Now we want to go for Naive Bayes approximation. What will the class-conditional distributions of X1 and X2 under label A, i.e. prob(X1 | Y=A) and prob (X2 | Y=A)? What will be the prediction and risk at the point (1,2) using this Naïve Bayes classifier?