Problem 3:

a)

P(x1=1|y=1) = 3/4
P(x1=1|y=-1=1/2
P(x1=0|y=1)=1/4
P(x1=0|y=-1)=1/2

P(x2=1|y=1)=0
P(x2=1|y=-1)=5/6
P(x2=0|y=1)= 4/4= 1
P(x2=0|y=-1)=1/6

P(x3=1|y=1)=3/4
P(x3=1|y=-1)=4/6
P(x3=0|y=1)=1/4
P(x3=0|y=-1)=2/6

P(x4=1|y=1)=2/4
P(x4=1|y=-1)=5/6
P(x4=0|y=1)=2/4
P(x4=0|y=-1)=1/6

P(x5=1|y=1)=1/4
P(x5=1|y=-1)=2/6
P(x5=0|y=1)=3/4
P(x5=0|y=-1)=4/6

b)

X= (0 0 0 0 0)
Naïve Bayesian rule: P(y|X) = P(x1|y)*P(x2|y)*P(x3|y)*P(x4|y)*P(x5|y)*P(y)
For y=1,
P(y=1| X= (0 0 0 0 0)) = (1/4)*1*(1/4)*(1/2)*(3/4) *(4/10)
= 0.009375

P(y=-1| X= (0 0 0 0 0))= (1/2)*(1/6)*(2/6)*(1/6)*(4/6) * (6/10)
= 0.001851

Class predicted for X= (0 0 0 0 0) is Y=1.

X= (1 1 0 1 0)
For y=1,
P(y=1| X= (1 1 0 1 0)) = (3/4)*(0)*…..
= 0

For y=-1,
P(y=-1| X= (1 1 0 1 0)) = (1/2)*(5/6)*(1/3)*(5/6)*(4/6)*(6/10)
= 0.04629

Class predicted for X= (1 1 0 1 0)) is Y= -1

c)
P(y=1|x=11010) = P(x=11010|y=1)*P(y=1)/P(x=11010)
= 0

d)

We should not use a Joint Bayes classifier, in contrast to Naïve Bayes classifier, because in Joint Bayes classifier, we shall have 5 features, each having 2 possible values, hence, 2^5=32 independent probabilities to estimate classification probability. Whereas in Naïve Bayes classifier, we have 4 and 6 data points. For both the cases, we have only 10 observations and thus, Bayes classifier using Joint probability would be uncertain to generalize well with new data.

e) There is no need to re-train the model, as in Naïve Bayes classifier, classification is done based on each feature individually, and the features are independent of each other.