a)
P(Y=1) is equal to 0.4
P (Xi=1|Y=1) [0.75 0.   0.75 0.5   0.25]
P(Xi=1|Y=-1) [0.5 0.83333333 0.66666667 0.83333333 0.33333333])

b)
for 0 0 0 0 0
p(x1=0,x2=0,x3=0,x4=0,x5=0|y=1)p(y=1) = 0.009375
p(x1=0,x2=0,x3=0,x4=0,x5=0|y=-1)p(y=-1) = 0.00185185185185
because 0.009375 is greater than 0.00185185185185 the predicted class i
s Y= 1 for 0 0 0 0 0

for 1 1 0 1 0
Because P(X2=1|Y=Read) is equal to zero and as its not mentioned to use s
moothing or adding an alpha,I leave it az 0
P(x1=1,x2=1,x3=0,x4=1,x5=0|y=1)p(y=1) = 0.0
P(x1=1,x2=1,x3=0,x4=1,x5=0|y=-1)p(y=-1) = 0.0462962962963
Y= -1 is predicted as it maximizes likelihood for 1 1 0 1 0

c)
As its not mentioned to smooth or add an alpha in case if a probablity of
 an independent variable for a given class
is Zero, the posterior probability P(Y=+1|x1=1,x2=1,x3=0,x4=1,x5=1) equal
s to Zero , because P(X2=1|Y=+1)= 0 and
then P(x1=1,x2=1,x3=0,x4=1,x5=0|y=1)p(Y=+1) = 0.0

d)
If we use joint Bayes classifier then our join distribution table would b
e order O((number of features) power 2))
as opposed to Naive Bayes Classifier which is linear and is order O(numb
er of features)as we assume that each variable is independent

e)
In Naive Bayes we assume that all vairables are independent ,
so being not able to tell whether we know the author or not , we still ca
n use our model
using other variables X2...X5 for prediction purpose