(1)

A = {000,001,010,011}
B = {000,001,100,101}
C = {000,100,010,001}

P(A) = 1/2
P(B) = 1/2
P(C) = 1/2
P(A&B) = 1/4
P(A&C) = 3/8
P(B&C) = 3/8

(2)

The Naive Bayes assumption of conditional independence is that we assume that the Elements are Independent to another even if there are not.
The advantage of this assumption is that we can build easier Classifiers because we don't have to understand the dependency of the words.

(3)

If we don't add a prior it could happen that the probability a text belonging to any class could be 0.

e.g.
We have 2 Classes and 3 Known words.
1st Class x={a|0, b|1, c|0}
2nd Class y= {a|0.5, b|0, c|0.5}

now if we want the best suited class for "a b" the probability of belonging to x and y would be 0.

if we would add a small prior we would get something like:
1st Class x={a|0.048, b|0.904, c|0.048}
2nd Class y= {a|0.476, b|0.048, c|0.476}

now Q=P(x)/P(y) would be 1,90