

1. A spam filter is designed by looking at commonly occurring phrases in spam. Suppose that 80% of email is spam. In 10% of the spam emails, the phrase “free money” is used, whereas this phrase is only used in 1% of non-spam emails. A new email has just arrived, which does mention “free money”. What is the probability that it is spam?

ANS:

A: The event email is spam

B: The event of email has free money

$$P(A|B) = P(B|A) * P(A)/P(B)$$

$$0.1/0.8/(0.1*0.8)+(0.01*0.2)$$

$$= 80/82$$