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?

Answer :

To find the probability that an email is spam given that it contains the phrase "free money," we can use **Bayes' Theorem**.

**Definitions and Given Information**

* **Let SSS** be the event that an email is spam.
* **Let NNN** be the event that an email is not spam.
* **Let FFF** be the event that the email contains the phrase "free money."

Given:

* P(S)=0.80P(S) = 0.80P(S)=0.80: Probability that an email is spam.
* P(N)=0.20P(N) = 0.20P(N)=0.20: Probability that an email is not spam.
* P(F∣S)=0.10P(F | S) = 0.10P(F∣S)=0.10: Probability that the phrase "free money" is used in a spam email.
* P(F∣N)=0.01P(F | N) = 0.01P(F∣N)=0.01: Probability that the phrase "free money" is used in a non-spam email.

**Applying Bayes' Theorem**

Bayes' Theorem states:

P(S∣F)=P(F∣S)×P(S)P(F)P(S | F) = \frac{P(F | S) \times P(S)}{P(F)}P(S∣F)=P(F)P(F∣S)×P(S)​

We need to calculate P(F)P(F)P(F), the total probability that the email contains the phrase "free money." This can be found using the Law of Total Probability:

P(F)=P(F∣S)×P(S)+P(F∣N)×P(N)P(F) = P(F | S) \times P(S) + P(F | N) \times P(N)P(F)=P(F∣S)×P(S)+P(F∣N)×P(N)

Substituting the values:

P(F)=(0.10×0.80)+(0.01×0.20)P(F) = (0.10 \times 0.80) + (0.01 \times 0.20)P(F)=(0.10×0.80)+(0.01×0.20) P(F)=0.08+0.002=0.082P(F) = 0.08 + 0.002 = 0.082P(F)=0.08+0.002=0.082

Now, applying Bayes' Theorem:

P(S∣F)=P(F∣S)×P(S)P(F)P(S | F) = \frac{P(F | S) \times P(S)}{P(F)}P(S∣F)=P(F)P(F∣S)×P(S)​ P(S∣F)=0.10×0.800.082P(S | F) = \frac{0.10 \times 0.80}{0.082}P(S∣F)=0.0820.10×0.80​ P(S∣F)=0.080.082≈0.9756P(S | F) = \frac{0.08}{0.082} \approx 0.9756P(S∣F)=0.0820.08​≈0.9756

**Final Answer**

The probability that the email is spam given that it contains the phrase "free money" is approximately **0.9756** or **97.56%**.