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?

Solution

To solve this problem, we can use \*\*Bayes’ Theorem\*\*, which relates the conditional probability of events.

Let’s define the events:

- \( S \): The event that an email is spam.

- \( N \): The event that an email is not spam (i.e., it is normal).

- \( F \): The event that the email contains the phrase “free money”.

We are given:

- \( P(S) = 0.8 \): The probability that any email is spam.

- \( P(N) = 0.2 \): The probability that any email is not spam (since \( P(N) = 1 – P(S) \)).

- \( P(F|S) = 0.1 \): The probability that an email contains “free money” given that it is spam.

- \( P(F|N) = 0.01 \): The probability that an email contains “free money” given that it is not spam.

We want to find \( P(S|F) \): The probability that an email is spam given that it contains “free money”.

### Bayes’ Theorem:

\[

P(S|F) = \frac{P(F|S) \cdot P(S)}{P(F)}

\]

Where \( P(F) \) is the total probability that an email contains “free money”. We can calculate \( P(F) \) using the law of total probability:

\[

P(F) = P(F|S) \cdot P(S) + P(F|N) \cdot P(N)

\]

Now, plug in the values:

\[

P(F) = (0.1 \times 0.8) + (0.01 \times 0.2)

\]

\[

P(F) = 0.08 + 0.002 = 0.082

\]

Finally, use Bayes’ Theorem to find \( P(S|F) \):

\[

P(S|F) = \frac{0.1 \times 0.8}{0.082} = \frac{0.08}{0.082} \approx 0.9756

\]

### Conclusion:

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