1 Lesson 9 Example 2

One application where Bayes' Theorem has been extremely successful is spam filtering. From historical data, 80% of all email is spam, and the phrase "free money" is used in 10% of spam emails. That is, $P(\text{"free money"} \mid \text{spam}) = 0.1$. The phrase is also used in 1% of non-spam emails. A new email has just arrived which contains the phrase "free money". Given this information, what is the probability that it is spam, $P(\text{spam} \mid \text{"free money"})$?

2 Answer

2.1 Given

- $P(\text{spam}) = 0.80$ (80% of all emails are spam)
- $P(\text{"free money"} \mid \text{spam}) = 0.10$ (10% of spam emails contain the phrase "free money")
- $P(\text{"free money"} \mid \text{not spam}) = 0.01$ (1% of non-spam emails contain the phrase "free money")

We need to find $P(\text{spam} \mid \text{"free money"})$.

2.2 Applying Bayes' Theorem

Bayes' Theorem is given by:

$$P(\text{spam} \mid \text{"free money"}) = \frac{P(\text{"free money"} \mid \text{spam}) \cdot P(\text{spam})}{P(\text{"free money"})}$$

We can calculate $P(\text{"free money"})$ using the Law of Total Probability, which accounts for both spam and non-spam emails:

$$P(\text{"free money"}) = (P(\text{"free money"} \mid \text{spam}) \cdot P(\text{spam})) + (P(\text{"free money"} \mid \text{not spam}) \cdot P(\text{not spam}))$$

Substituting the known values:

$$P(\text{"free money"}) = (0.10 \times 0.80) + (0.01 \times 0.20)$$

$$P(\text{"free money"}) = 0.08 + 0.002 = 0.082$$

Now, using Bayes' Theorem:

$$P(\text{spam} \mid \text{"free money"}) = \frac{0.10 \times 0.80}{0.082}$$

$$P(\text{spam} \mid \text{"free money"}) = \frac{0.08}{0.082} \approx 0.9756$$

2.3 Conclusion

The probability that an email containing the phrase ”free money” is spam is approximately **0.9756**, or **97.56%**.