Solutions

Problem 1

(a) **Compute the Prior Probabilities**

The prior probabilities are computed as:

$$P(Y = 1) = \frac{\text{Number of patients with disease}}{\text{Total number of patients}}$$

$$P(Y = 0) = \frac{\text{Number of healthy patients}}{\text{Total number of patients}}$$

From the dataset:

$$P(Y = 1) = \frac{3}{8} = 0.375, \quad P(Y = 0) = \frac{5}{8} = 0.625$$

(b) **Compute the Conditional Probabilities**

The conditional probabilities are given by:

$$P(X_j = 1 \mid Y) = \frac{\text{Number of cases where } X_j = 1 \text{ in class } Y}{\text{Total cases in class } Y}$$

For $Y = 1$ (Has Disease):

$$P(X_1 = 1 \mid Y = 1) = \frac{2}{3} \approx 0.6667,$$
$$P(X_2 = 1 \mid Y = 1) = \frac{2}{3} \approx 0.6667,$$
$$P(X_3 = 1 \mid Y = 1) = \frac{2}{3} \approx 0.6667$$

For $Y = 0$ (Healthy):

$$P(X_1 = 1 \mid Y = 0) = \frac{2}{5} = 0.4000,$$
$$P(X_2 = 1 \mid Y = 0) = \frac{2}{5} = 0.4000,$$
$$P(X_3 = 1 \mid Y = 0) = \frac{2}{5} = 0.4000$$

(c) **Compute the Posterior Probabilities**

Since:

$$P(X_2 = 0 \mid Y) = 1 - P(X_2 = 1 \mid Y)$$

$$P(X_2 = 0 \mid Y = 1) = 1 - \frac{2}{3} = \frac{1}{3}, \quad P(X_2 = 0 \mid Y = 0) = 1 - \frac{2}{5} = \frac{3}{5}$$

Compute Numerator Values (Unnormalized Probabilities):

For $Y = 1$:

$$P(Y=1)P(X_1=1 \mid Y=1)P(X_2=0 \mid Y=1)P(X_3=1 \mid Y=1)$$

$$= \frac{3}{8} \times \frac{2}{3} \times \frac{1}{3} \times \frac{2}{3} = \frac{1}{18} \approx 0.056$$

For $Y=0$:

$$P(Y=0)P(X_1=1 \mid Y=0)P(X_2=0 \mid Y=0)P(X_3=1 \mid Y=0)$$

$$= \frac{5}{8} \times \frac{2}{5} \times \frac{3}{5} \times \frac{2}{5} = \frac{3}{50} = 0.06$$

Normalization Step:

The final posterior probabilities are obtained by dividing by the sum of both numerators:

$$P(Y=1 \mid X) = \frac{\frac{1}{18}}{\frac{1}{18} + \frac{3}{50}} = \frac{25}{52} \approx 0.481$$

$$P(Y=0 \mid X) = \frac{\frac{3}{50}}{\frac{1}{18} + \frac{3}{50}} = \frac{27}{52} \approx 0.519$$

Problem 2

Please refer to the linked .ipynb file.
https://drive.google.com/file/d/17aeDQQdXhJws7ycAzRkcqU99LCoQByGd/view?usp=sharing