ECE/CS/ME 539 – Fall 2024 — Activity 5

1.

Consider a bunch of eggs. The length (cm) of an egg is used as a feature to classify the eggs into either Large (L) or Jumbo (J). The distributions of eggs in each category (length and label) are as follows:

Label\Length (cm)	1	2	3	4	5
Large	3	8	6	2	0
Jumbo	0	1	5	9	6

(a) Total number of eggs in this bunch, $N = 40$

(b) Class distribution: $\Pr\{\text{Large}\} = 47.5\ \%$, $\Pr\{\text{Jumbo}\} = 52.5\ \%$

(c) Likelihood that the length of a large egg is longer than 3 cm: $\Pr\{x > 3 \mid \text{Large}\} = 10.53\ \%$

(d) The posterior probability that an egg is a jumbo egg given that its length is shorter than 4 cm: $\Pr\{\text{Jumbo} \mid x < 4\} = \frac{6}{23} = 26.09\ \%$

(e) $\Pr\{x < 4\} = \frac{23}{40} = 57.5\%$

(f) $\Pr\{\text{Jumbo}, x < 4\} = 15\ \%$
$\Pr\{\text{Jumbo}, x < 4\} = \Pr\{\text{Jumbo}, \quad \text{x} = 1\} + \Pr\{\text{Jumbo}, \quad \text{x} = 2\} + \Pr\{\text{Jumbo}, \quad \text{x} = 3\}$
$= \frac{0}{40} + \frac{1}{40} + \frac{5}{40} = \frac{6}{40} = 15\%.$

(g) Find the likelihood $\Pr\{x < 4 \mid \text{Jumbo}\} = 28.57\ \%$
$\text{Pr.}\{x < 4 \mid \text{Jumbo}\} = \frac{0+1+5}{0+1+5+9+6} = \frac{6}{21} = 28.57\%$

(h) Using the above results, verify the Bayesian equation

$$\Pr\{\text{Jumbo} \mid x < 4\} = \frac{\Pr\{x < 4 \mid \text{Jumbo}\} \cdot \Pr\{\text{Jumbo}\}}{\Pr\{x < 4\}}$$

$$\frac{\left(\frac{6}{21}\right) \cdot \left(\frac{21}{40}\right)}{\frac{23}{40}} = \frac{6}{23} = 26.09\%$$

(i) Consider the Bayesian decision rule (a.k.a. Maximum a Posterior, MAP decision rule)

$$g(x) = \begin{cases} \text{Large} & \text{if } \Pr\{\text{Large} \mid x\} > \Pr\{\text{Jumbo} \mid x\} \\ \text{Jumbo} & \text{otherwise} \end{cases}$$

Evaluate $g(x)$ for $x \in \{1, 2, 3, 4, 5\}$ (unit: cm)

Answer:

$$\Pr\{\text{Large} \mid 1\text{cm}\} = 1, \quad \Pr\{\text{Jumbo} \mid 1\text{cm}\} = 0. \text{ Hence } g(1\text{cm}) = \text{Large}$$

$$\Pr\{\text{Large} \mid 2\text{cm}\} = \frac{8}{9}, \quad \Pr\{\text{Jumbo} \mid 2\text{cm}\} = \frac{1}{8}. \text{ Hence } g(2\text{cm}) = \text{Large}$$

$$\Pr\{\text{Large} \mid 3\text{cm}\} = \frac{6}{11}, \quad \Pr\{\text{Jumbo} \mid 3\text{cm}\} = \frac{5}{11}. \text{ Hence } g(3\text{cm}) = \text{Large}$$

$$\Pr\{\text{Large} \mid 4\text{cm}\} = \frac{2}{11}, \quad \Pr\{\text{Jumbo} \mid 4\text{cm}\} = \frac{9}{11}. \text{ Hence } g(4\text{cm}) = \text{Jumbo}$$

$$\Pr\{\text{Large} \mid 5\text{cm}\} = 0, \quad \Pr\{\text{Jumbo} \mid 5\text{cm}\} = 1. \text{ Hence } g(5\text{cm}) = \text{Jumbo}$$

(j) Using the classifier developed in part (i), evaluate the probability of correct classification:
Answer:

For $x \leq 3$cm, $3 + 8 + 6 = 17$ large eggs are correctly labeled as Large, $0 + 1 + 5 = 6$ jumbo eggs are mistakenly labeled as Large.

For $x > 3$cm, $2 + 0 = 2$ Large eggs are mistakenly labeled as Jumbo eggs, and $9 + 6 = 15$ Jumbo eggs are correctly labeled as Jumbo.

Thus, the total number of correctly labeled (predicted) eggs $= 17 + 15 = 32$.

The probability of correct classification (classification rate) $= \frac{32}{40} = 80\%$.

2.

Table 1 lists the weather outlook, temperature, humidity, and wind conditions of 14 days and the outcome of whether a junior league will play footfall or not. Calculate the following empirical probabilities based on given weather conditions and junior league play outcomes:

(a) Pr.{Humidity = High} $= \frac{7}{14} = 0.5$

(b) Pr.{Outlook = Sunny AND Humidity = Normal} $= \frac{2}{14} = 14.29\%$

(c) Pr.{Temperature = Cool OR Wind = Weak} $= \frac{10}{14} = 71.4\%$

(d) Pr.{Play = Yes | Humidity = High} $= \frac{3}{7} = 42.86\%$

(e) Pr.{Humidity = High | Play = Yes} $= \frac{3}{9} = 33.33\%$

(f) Entropy of Play:

$$\text{Entropy(Play)} = -\left(\frac{9}{14}\log_2\frac{9}{14} + \frac{5}{14}\log_2\frac{5}{14}\right) = 0.9403 \text{ bit}$$

Outlook	**Temperature**	**Humidity**	**Wind**	**Played football (yes/no)**
Sunny	Hot	High	Weak	No
Sunny	Hot	High	Strong	No
Overcast	Hot	High	Weak	Yes
Rain	Mild	High	Weak	Yes
Rain	Cool	Normal	Weak	Yes
Rain	Cool	Normal	Strong	No
Overcast	Cool	Normal	Strong	Yes
Sunny	Mild	High	Weak	No
Sunny	Cool	Normal	Weak	Yes
Rain	Mild	Normal	Weak	Yes
Sunny	Mild	Normal	Strong	Yes
Overcast	Mild	High	Strong	Yes
Overcast	Hot	Normal	Weak	Yes
Rain	Mild	High	Strong	No

Table 1: Weather conditions and football playing decision