DATA622 - HW-1, Q-1

1) Compute Prior probabilities for the Prospect Yes/No

Answer:

$$P(\text{Prospect} = \text{yes}) = \frac{\text{number of prospect = 'Yes' in the original data}}{\text{number of records}}$$

$$= \frac{9}{14}$$

$$P(\text{Prospect} = \text{No}) = \frac{\text{number of prospect = 'No' in the original data}}{\text{Total number of records}}$$

$$= \frac{5}{14}$$

---

2) Conditional probabilities:

We will use the definition formula for conditional probability.

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$(a)\ P(\text{age-group} = \text{youth} \mid \text{prospect} = \text{yes}) = \frac{P(\text{age-group} = \text{youth} \ \& \ \text{prospect} = \text{yes})}{P(\text{prospect} = \text{yes})}$$

using/ taking the counts

$$= \frac{\dfrac{\text{number of records with age-grp = youth \& prospect = yes}}{\text{Total number of records}}}{\dfrac{\text{number of records with prospect = yes}}{\text{Total number of records}}}$$

$$= \frac{2/14}{9/14} = \frac{2}{9}$$

② b) $P(\text{age-group} = \text{middle} \mid \text{prospect} = \text{yes}) = \dfrac{P(\text{agegroup} = \text{middle and prospect} = \text{yes})}{P(\text{prospect} = \text{yes})}$

$$= \frac{4}{9}$$

c) $P(\text{age-group} = \text{senior} \mid \text{prospect} = \text{yes}) = \dfrac{P(\text{age-group} = \text{senior} \;\&\; \text{prospect} = \text{yes})}{P(\text{prospect} = \text{yes})}$

$$= \frac{3}{9} = \frac{1}{3}$$

d) $P(\text{age-group} = \text{youth} \mid \text{prospect} = \text{no}) = \dfrac{P(\text{age-group} = \text{youth} \;\&\; \text{prospect} = \text{no})}{P(\text{prospect} = \text{no})}$

$$= \frac{3/14}{5/14} = \frac{3}{5}$$

e) $P(\text{age-group} = \text{middle} \mid \text{prospect} = \text{no}) = \dfrac{P(\text{age-group} = \text{middle} \;\&\; \text{prospect} = \text{no})}{P(\text{prospect} = \text{no})}$

$$= \frac{0/14}{5/14} = 0$$

f) $P(\text{age-group} = \text{senior} \mid \text{prospect} = \text{no}) = \dfrac{P(\text{age-group} = \text{senior} \;\&\; \text{prospect} = \text{no})}{P(\text{prospect} = \text{no})}$

$$= \frac{2/14}{5/14} = \frac{2}{5}$$

g) $P(\text{networth} = \text{high} \mid \text{prospect} = \text{yes}) = \dfrac{P(\text{networth} = \text{high} \;\&\; \text{prospect} = \text{yes})}{P(\text{prospect} = \text{yes})}$

$$= \frac{2/14}{9/14} = \frac{2}{9}$$

h) $P(\text{networth} = \text{medium} \mid \text{prospect} = \text{yes}) = \dfrac{P(\text{networth} = \text{medium} \;\&\; \text{prospect} = \text{yes})}{P(\text{prospect} = \text{yes})}$

$$= \frac{4/14}{9/14} = \frac{4}{9}$$

③

i) $P(\text{networth} = \text{low} \mid \text{prospect} = \text{yes}) = \dfrac{P(\text{networth} = \text{low} \ \& \ \text{prospect} = \text{yes})}{P(\text{prospect} = \text{yes})}$

$= \dfrac{3/14}{9/14} = \dfrac{3}{9} = \dfrac{1}{3}$

j) $P(\text{networth} = \text{high} \mid \text{prospect} = \text{no}) = \dfrac{P(\text{networth} = \text{high} \ \& \ \text{prospect} = \text{no})}{P(\text{prospect} = \text{no})}$

$= \dfrac{2/14}{5/14} = \dfrac{2}{5}$

k) $P(\text{networth} = \text{medium} \mid \text{prospect} = \text{no}) = \dfrac{P(\text{networth} = \text{medium} \ \& \ \text{prospect} = \text{no})}{P(\text{prospect} = \text{no})}$

$= \dfrac{2/14}{5/14} = \dfrac{2}{5}$

l) $P(\text{networth} = \text{low} \mid \text{prospect} = \text{no}) = \dfrac{P(\text{networth} = \text{low} \ \& \ \text{prospect} = \text{no})}{P(\text{prospect} = \text{no})}$

$= \dfrac{1/14}{5/14} = \dfrac{1}{5}$

m) $P(\text{status} = \text{employed} \mid \text{prospect} = \text{yes}) = \dfrac{P(\text{status} = \text{employed} \ \& \ \text{prospect} = \text{yes})}{P(\text{prospect} = \text{yes})}$

$= \dfrac{3/14}{9/14} = \dfrac{3}{9} = \dfrac{1}{3}$

n) $P(\text{status} = \text{unemployed} \mid \text{prospect} = \text{yes}) = \dfrac{P(\text{status} = \text{unemployed} \ \& \ \text{prospect} = \text{yes})}{P(\text{prospect} = \text{yes})}$

$= \dfrac{6/14}{9/14} = \dfrac{6}{9} = \dfrac{2}{3}$

o) $P(\text{status} = \text{employed} \mid \text{prospect} = \text{no}) = \dfrac{P(\text{status} = \text{employed} \ \& \ \text{prospect} = \text{no})}{P(\text{prospect} = \text{no})}$

$= \dfrac{4/14}{5/14} = \dfrac{4}{5}$

p) $P(\text{status} = \text{unemployed} \mid \text{prospect} = \text{no}) = \dfrac{P(\text{status} = \text{unemployed} \ \& \ \text{prospect} = \text{no})}{P(\text{prospect} = \text{no})}$

$= \dfrac{1/14}{5/14} = \dfrac{1}{5}$

(4) q) $P(\text{credit} = \text{fair} \mid \text{prospect} = \text{yes}) = \frac{P(\text{credit} = \text{fair} \ \& \ \text{prospect} = \text{yes})}{P(\text{prospect} = \text{yes})}$

$$= \frac{6/14}{9/14} = \frac{6}{9} = \frac{2}{3}$$

r) $P(\text{credit} = \text{excellent} \mid \text{prospect} = \text{yes}) = \frac{P(\text{credit} = \text{excellent} \ \& \ \text{prospect} = \text{yes})}{P(\text{prospect} = \text{yes})}$

$$= \frac{3/14}{9/14} = \frac{3}{9} = \frac{1}{3}$$

s) $P(\text{credit} = \text{fair} \mid \text{prospect} = \text{no}) = \frac{P(\text{credit} = \text{fair} \ \& \ \text{prospect} = \text{no})}{P(\text{prospect} = \text{no})}$

$$= \frac{2/14}{5/14} = \frac{2}{5}$$

t) $P(\text{credit} = \text{excellent} \mid \text{prospect} = \text{no}) = \frac{P(\text{credit} = \text{excellent} \ \& \ \text{prospect} = \text{no})}{P(\text{prospect} = \text{no})}$

$$= \frac{3/14}{5/14} = \frac{3}{5}$$

---

3) Posterior probabilities

Before we calculate Posterior probabilities, let us calculate the prior probabilities of each categorical feature other than prospect - which we already calculated in (1).

$P(\text{age-group} = \text{youth}) = \frac{5}{14}$

$P(\text{age-group} = \text{middle}) = \frac{4}{14}$

$P(\text{age-group} = \text{senior}) = \frac{5}{14}$

$P(\text{networth} = \text{high}) = \frac{4}{14}$

$P(\text{networth} = \text{medium}) = \frac{6}{14}$

$P(\text{networth} = \text{low}) = \frac{4}{14}$

$P(\text{status} = \text{employed}) = \frac{7}{14}$

$P(\text{status} = \text{unemployed}) = \frac{7}{14}$

$P(\text{credit-rating} = \text{fair}) = \frac{8}{14}$

$P(\text{credit-rating} = \text{excellent}) = \frac{6}{14}$

⑤ Posterior probabilities

a) $P(\text{prospect} = \text{no} \mid \text{age-group} = \text{youth})$.

Using Bayes Theorem,

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B)}$$

Hence, $P(\text{prospect} = \text{no} \mid \text{age-group} = \text{youth}) = \frac{P(\text{age-group} = \text{youth} \mid \text{prospect} = \text{no}) P(\text{prospect} = \text{no})}{P(\text{age-group} = \text{youth})}$

$$= \frac{\frac{3}{5} \times \frac{5}{14}}{\frac{5}{14}} = \frac{3}{5}$$

b) $P(\text{prospect} = \text{no} \mid \text{age-group} = \text{middle}) = \frac{P(\text{age-group} = \text{middle} \mid \text{prospect} = \text{no}) P(\text{prospect} = \text{no})}{P(\text{age-group} = \text{middle})}$

$$= \frac{0 \times \frac{5}{14}}{4/14} = 0$$

c) $P(\text{prospect} = \text{no} \mid \text{age-group} = \text{senior}) = \frac{P(\text{age-group} = \text{senior} \mid \text{prospect} = \text{no}) P(\text{prospect} = \text{no})}{P(\text{age-group} = \text{senior})}$

$$= \frac{\frac{2}{5} \times \frac{5}{14}}{\frac{5}{14}} = \frac{2}{5}$$

d) $P(\text{prospect} = \text{no} \mid \text{networth} = \text{high}) = \frac{P(\text{networth} = \text{high} \mid \text{prospect} = \text{no}) P(\text{prospect} = \text{no})}{P(\text{networth} = \text{high})}$

$$= \frac{\frac{2}{5} \times \frac{5}{14}}{\frac{4}{14}} = \frac{2}{4} = \frac{1}{2}$$

e) $P(\text{prospect} = \text{no} \mid \text{networth} = \text{medium}) = \frac{P(\text{networth} = \text{medium} \mid \text{prospect} = \text{no}) P(\text{prospect} = \text{no})}{P(\text{networth} = \text{~~high~~ medium})}$

$$= \frac{\frac{2}{5} \times \frac{5}{14}}{\frac{6}{14}} = \frac{2}{6} = \frac{1}{3}$$

f) $P(\text{prospect} = \text{no} \mid \text{networth} = \text{low}) = \frac{P(\text{networth} = \text{low} \mid \text{prospect} = \text{no}) P(\text{prospect} = \text{no})}{P(\text{networth} = \text{low})}$

$$= \frac{\frac{1}{5} \times \frac{5}{14}}{4/14} = \frac{1}{4}$$

(6) g) $P(\text{prospect}=\text{no} \mid \text{status}=\text{employed}) = \dfrac{P(\text{status}=\text{employed} \mid \text{prospect}=\text{no})\, P(\text{prospect}=\text{no})}{P(\text{status}=\text{employed})}$

$= \dfrac{4/5 \times 5/14}{7/14} = \dfrac{4}{7}$

h) $P(\text{prospect}=\text{no} \mid \text{status}=\text{unemployed}) = \dfrac{P(\text{status}=\text{unemployed} \mid \text{prospect}=\text{no})\, P(\text{prospect}=\text{no})}{P(\text{status}=\text{unemployed})}$

$= \dfrac{1/5 \times 5/14}{7/14} = \dfrac{1}{7}$

i) $P(\text{prospect}=\text{no} \mid \text{credit}=\text{fair}) = \dfrac{P(\text{credit}=\text{fair} \mid \text{prospect}=\text{no})\, P(\text{prospect}=\text{no})}{P(\text{credit}=\text{fair})}$

$= \dfrac{2/5 \times 5/14}{8/14} = \dfrac{1}{4}$

j) $P(\text{prospect}=\text{no} \mid \text{credit}=\text{excellent}) = \dfrac{P(\text{credit}=\text{excellent} \mid \text{prospect}=\text{no})\, P(\text{prospect}=\text{no})}{P(\text{credit}=\text{excellent})}$

$= \dfrac{3/5 \times 5/14}{6/14} = \dfrac{3}{6} = \dfrac{1}{2}$

k) $P(\text{prospect}=\text{yes} \mid \text{age-group}=\text{youth}) = \dfrac{P(\text{age-group}=\text{youth} \mid \text{prospect}=\text{yes})\, P(\text{prospect}=\text{yes})}{P(\text{age-group}=\text{youth})}$

$= \dfrac{2/9 \times 9/14}{5/14} = \dfrac{2}{5}$

l) $P(\text{prospect}=\text{yes} \mid \text{age-group}=\text{middle}) = \dfrac{P(\text{age-group}=\text{middle} \mid \text{prospect}=\text{yes})\, P(\text{prospect}=\text{yes})}{P(\text{age-group}=\text{middle})}$

$= \dfrac{4/9 \times 9/14}{4/14} = 1$

m) $P(\text{prospect}=\text{yes} \mid \text{age-group}=\text{Senior}) = \dfrac{P(\text{age-group}=\text{Senior} \mid \text{prospect}=\text{yes})\, P(\text{prospect}=\text{yes})}{P(\text{age-group}=\text{Senior})}$

$= \dfrac{1/3 \times 9/14}{5/14} = \dfrac{3}{5}$

n) $P(\text{prospect}=\text{yes} \mid \text{networth}=\text{high}) = \dfrac{P(\text{networth}=\text{high} \mid \text{prospect}=\text{yes})\, P(\text{prospect}=\text{yes})}{P(\text{networth}=\text{high})}$

$= \dfrac{2/9 \times 9/14}{4/14} = \dfrac{1}{2}$

7. o) $P(\text{prospect=yes} \mid \text{networth = medium}) = \frac{P(\text{networth = medium} \mid \text{prospect=yes})\, P(\text{prospect=yes})}{P(\text{networth = medium})}$

$$= \frac{\frac{4}{9} \times \frac{9}{14}}{\frac{6}{14}} = \frac{2}{3}$$

p) $P(\text{prospect=yes} \mid \text{networth = low}) = \frac{P(\text{networth = low} \mid \text{prospect = yes})\, P(\text{prospect = yes})}{P(\text{networth = low})}$

$$= \frac{\frac{1}{3} \times \frac{9}{14}}{\frac{4}{14}} = \frac{3}{4}$$

q) $P(\text{prospect = yes} \mid \text{status = employed}) = \frac{P(\text{status = employed} \mid \text{prospect = yes})\, P(\text{prospect = yes})}{P(\text{status = employed})}$

$$= \frac{\frac{1}{3} \times \frac{9}{14}}{\frac{7}{14}} = \frac{3}{7}$$

r) $P(\text{prospect = yes} \mid \text{status = unemployed}) = \frac{P(\text{status = unemployed} \mid \text{prospect = yes})\, P(\text{prospect = yes})}{P(\text{status = unemployed})}$

$$= \frac{\frac{2}{3} \times \frac{9}{14}}{\frac{7}{14}} = \frac{6}{7}$$

s) $P(\text{prospect=yes} \mid \text{credit = fair}) = \frac{P(\text{credit = fair} \mid \text{prospect = yes})\, P(\text{prospect = yes})}{P(\text{credit = fair})}$

$$= \frac{\frac{2}{3} \times \frac{9}{14}}{\frac{8}{14}} = \frac{3}{4}$$

t) $P(\text{prospect=yes} \mid \text{Credit = excellent}) = \frac{P(\text{credit = excellent} \mid \text{prospect=yes})\, P(\text{prospect=yes})}{P(\text{Credit = excellent})}$

$$= \frac{\frac{1}{3} \times \frac{9}{14}}{\frac{6}{14}} = \frac{3}{6} = \frac{1}{2}$$

P