Name Kumudini Bhave

Title Data 622 Homework 1
Question 1

Given is the prospecting dataset

1. Compute prior probabilities for Yes / No.

P (Prospect = Yes)

= Total no of prospects with Yes / Total no of observations

= 9/14

P (prospect = No)

= Total no. of prospects with No / Total no. of observations

= 5/14

2. Conditional Probabilites.

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

2.1 $P(\text{age group} = \text{youth} \mid \text{prospect} = \text{yes})$

$= 2/9 = 0.22$

2.2 $P(\text{age group} = \text{middle} \mid \text{prospect} = \text{yes})$

$= 4/9 = 0.44$

2.3 $P(\text{age group} = \text{senior} \mid \text{prospect} = \text{yes})$

$= 3/9 = 0.33$

2.4 $P(\text{age-group} = \text{youth} \mid \text{prospect} = \text{no})$

$= 3/5 = 0.6$

2.5 $P(\text{age group} = \text{middle} \mid \text{prospect} = \text{no})$

$= 0/5 = 0$

2.6 $P(\text{age group} = \text{senior} \mid \text{prospect} = \text{no})$

$= 2/5 = 0.4$

2.7 P(networth = high | prospect = yes)

= 2/9 = 0.22

2.8 P(networth = medium | prospect = yes)

= 4/9 = 0.44

2.9 P(networth = low | prospect = yes)

= 3/9 = 0.33

2.10 P(networth = high | prospect = no)

= 2/5 = ~~0.40~~ → 0.40

2.11 P(networth = medium | prospect = no)

= 2/5 = 0.4

2.12 P(networth = low | prospect = no)

= 1/5 = 0.2

2.13 P(status = employed | prospect = yes)
= 3/9 = 0.333

2.14 P(status = unemployed | prospect = yes)
= 6/9 = 0.666

2.15 P(status = employed | prospect = no)
= 4/5 = 0.8

2.16 P(status = unemployed | prospect = no)
= 1/5 = 0.2

2.17 P(credit-rating = fair | prospect = yes)
= 6/9 = 0.666

2.18 P(credit-rating = excellent | prospect = yes)
= 3/9 = 0.33

2.19 P(credit rating = fair | prospect = no)
= 2/5 = 0.4

2.20 $P(\text{credit rating} = \text{excellent} \mid \text{prospect} = \text{no})$

$= 3/5 = 0.6$

3. Posterior Probabilities

Naive Bayes Theorem

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

Applying the above for each class of prospect for different predictors

3.1 $P(\text{prospect} = \text{yes} \mid \text{age-group} = \text{youth})$

$$= \frac{2/4 \cdot 9/14}{5/14} = \frac{2}{5} = 0.4$$

3.2 P(prospect = No | age group = youth)

$$\frac{3/5 \cdot 5/14}{5/14} = 3/5 = 0.6$$

3.3 P(prospect = Yes | age-group = middle)

$$\frac{4/9 \cdot 9/14}{4/14} = 4/4 = 1$$

3.4 P(prospect = No | age group = middle)

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

3.5 P(prospect = yes | age-group = senior)

$$\frac{3/9 \cdot 9/14}{5/14} = 3/5 = 0.6$$

3.6 P(prospect = No | age-group = senior)

$$\frac{2/5 \cdot 5/14}{5/14} = 2/5 = 0.4$$

3.7 $P(\text{prospect} = \text{yes} \mid \text{networth} = \text{high})$

$$\frac{2/9 \quad 9/14}{4/14} = 2/4 = 1/2 = 0.5$$

3.8 $P(\text{prospect} = \text{No} \mid \text{networth} = \text{high})$

$$\frac{2/5 \quad 5/14}{4/14} = 2/4 = 1/2 = 0.5$$

3.9 $P(\text{prospect} = \text{yes} \mid \text{networth} = \text{medium})$

$$\frac{4/9 \quad 9/14}{6/14} = 4/6 = 2/3 = 0.66$$

3.10 $P(\text{prospect} = \text{No} \mid \text{networth} = \text{medium})$

$$\frac{2/5 \quad 5/14}{6/14} = 2/6 = 1/3 = 0.33$$

3.11 $P(\text{prospect} = \text{yes} \mid \text{networth} = \text{low})$

$$\frac{3/9 \quad 9/14}{4/14} = \frac{1/3 \cdot 9/14}{4/14} = 3/4 = 0.75$$

3.12 $P(\text{prospect} = \text{No} \mid \text{networth} = \text{low})$

$$\frac{1/5 \quad 5/14}{4/14} \qquad = 1/4 = 0.25$$

3.13 $P(\text{prospect} = \text{yes} \mid \text{status} = \text{employed})$

$$3/9 \quad {}^{9}/_{14} \;\cdot\; {}^{1}/_{14} = \quad 3/7 = 0.428$$

3.14 $P(\text{prospect} = \text{No} \mid \text{status} = \text{employed})$

$$\frac{4/5 \quad 5/14}{7/14} = 4/7 = 0.571$$

3.15 $P(\text{prospect} = \text{yes} \mid \text{status} = \text{unemployed})$

$$\frac{6/9 \cdot 9/14}{7/14} \qquad \frac{2/3 \cdot 9/14}{7/14} = 6/7 \qquad \overset{0.857}{2 0.857}$$

3.16 $P(\text{prospect} = \text{No} \mid \text{status} = \text{unemployed})$

$$\frac{1/5 \quad 5/14}{7/14} = 1/7 = 0.142.$$

3.17 $P(\text{prospect} = \text{yes} \mid \text{credit-rating} = \text{fair})$

$$\frac{6/9 \quad 9/14}{8/14} = \frac{2/3 \cdot 9/14}{8/14} = 3/4 = 0.75$$

3.18 $P(\text{prospect} = \text{No} \mid \text{credit rating} = \text{fair})$

$$\frac{2/5 \quad 5/14}{8/14} = 1/4 = 0.25$$

3.19 $P(\text{prospect} = \text{yes} \mid \text{credit rating} = \text{excellent})$

$$\frac{3/9 \cdot 9/14}{6/14} = 1/2 = 0.5$$

3.20 $P(\text{prospect} = \text{No} \mid \text{credit-rating} = \text{excellent})$

$$\frac{3/5 \cdot 5/14}{6/14} = 1/2 = 0.5$$