Reading

4-4 Naive Bayes Classifier pp. 212–227,

Textbook Problems

p. 348

6a (6 pts), 6b (2 pts), 6c (2 pts), 6d (10 pts)

Hint: 6a Let: U-undergraduate, G-graduate, S-smoker. Then $P(S|U) = 0.15$, etc.

7a (6 pts), 7b (10 pts)

8a (6 pts), 8b (10 pts), 8c (2 pts), 8d (2 pts), 8e (2 pts)

6. a)

	smoke	all	dorm
undergrad	0.15	$\frac{4}{5}$	0.1
grad	0.23	$\frac{1}{5}$	0.3

$$P(\text{grad}|\text{smoke}) = \frac{P(\text{smoke}|\text{grad})P(\text{grad})}{P(\text{smoke})}$$

$P(\text{smoke}) \to P(S|U)P(U) + P(S|G)P(G)$

$$= \frac{0.23 \cdot \frac{1}{5}}{0.15 \cdot \frac{4}{5} + 0.23 \cdot \frac{1}{5}}$$

$$= 0.2771$$

b) $\because \frac{4}{5} > \frac{1}{5}$ $\therefore$ He/she is more likely to be undergraduate student

c) $\because P(U|S) > P(G|S)$ $\therefore$ He/she's more likely to be undergraduate student

$0.7229 > 0.2771$

d) $$P(\text{grad}|\text{smoke in dorm}) = \frac{P(\text{smoke in dorm}|\text{grad})P(\text{grad})}{P(\text{smoke in dorm})}$$

$P(\text{smoke in dorm}|\text{grad}) \to P(S|G)P(D|G)$; $P(\text{grad}) \to \frac{1}{5}$

$P(\text{smoke in dorm}) \to P(S|G)P(D|G)P(G) + P(S|U)P(D|U)P(U)$

$$= \frac{0.23 \cdot 0.3 \cdot \frac{1}{5}}{0.23 \cdot \frac{1}{5} \cdot 0.3 + 0.15 \cdot \frac{4}{5} \cdot 0.1}$$

$$= \frac{0.0138}{0.0258} \approx 0.5349$$

$P(\text{undergrad} \mid \text{smoke}) = 1 - 0.5349 = 0.4651 < P(\text{grad} \mid \text{smoke})$

$\therefore$ He/she's more likely to be a graduate student

7\. a) conditioned independence: $P(A \cap B \mid C) = P(A \mid C)\, P(B \mid C)$

$P(A=0 \mid +) = \frac{2}{5}$ $\quad$ $P(A=0 \mid -) = \frac{3}{5}$

$P(A=1 \mid +) = \frac{3}{5}$ $\quad$ $P(A=1 \mid -) = \frac{2}{5}$

$P(B=0 \mid +) = \frac{4}{5}$ $\quad$ $P(B=0 \mid -) = \frac{3}{5}$

$P(B=1 \mid +) = \frac{1}{5}$ $\quad$ $P(B=1 \mid -) = \frac{2}{5}$

$P(C=0 \mid +) = \frac{1}{5}$ $\quad$ $P(C=0 \mid -) = \frac{0}{5}$

$P(C=1 \mid +) = \frac{4}{5}$ $\quad$ $P(C=1 \mid -) = \frac{5}{5}$

b) $$P(+ \mid A=0, B=1, C=0) = \frac{P(A=0, B=1, C=0 \mid +)P(+)}{P(A=0, B=1, C=0)}$$

$$= \frac{P(A=0 \mid +)\, P(B=1 \mid +)\, P(C=0 \mid +)\, P(+)}{P(A=0, B=1, C=0 \mid +)P(+) + P(A=0, B=1, C=0 \mid -)P(-)}$$

$$= \frac{\frac{2}{5} \cdot \frac{1}{5} \cdot \frac{1}{5} \cdot \frac{5}{10}}{\frac{2}{5} \cdot \frac{1}{5} \cdot \frac{1}{5} \cdot \frac{5}{10} + \frac{3}{5} \cdot \frac{2}{5} \cdot \frac{0}{5} \cdot \frac{5}{10}} = 1$$

$$P(- \mid A=0, B=1, C=0) = \frac{P(A=0, B=1, C=0 \mid -)P(-)}{P(A=0, B=1, C=0)}$$

$$= \frac{P(A=0, B=1, C=0 \mid -)\, P(-)}{P(A=0, B=1, C=0 \mid -)\, P(-) + P(A=0, B=1, C=0 \mid +)\, P(+)}$$

$$= \frac{\frac{3}{5} \cdot \frac{2}{5} \cdot \frac{0}{5} \cdot \frac{5}{10}}{\frac{3}{5} \cdot \frac{2}{5} \cdot \frac{0}{5} \cdot \frac{5}{10} + \frac{2}{5} \cdot \frac{1}{5} \cdot \frac{1}{5} \cdot \frac{5}{10}} = 0$$

8. a) $P(A=1|+) = \frac{3}{5}$ $\quad P(B=1|+) = \frac{2}{5}$ $\quad P(C=1|+) = \frac{4}{5}$

$P(A=1|-) = \frac{2}{5}$ $\quad P(B=1|-) = \frac{2}{5}$ $\quad P(C=1|-) = \frac{1}{5}$

b) $$P(+|A=1, B=1, C=1) = \frac{P(A=1, B=1, C=1|+)\,P(+)}{P(A=1, B=1, C=1|+)\,P(+) \quad P(A=1, B=1, C=1|-)\,P(-)}$$

$$= \frac{\frac{3}{5} \cdot \frac{2}{5} \cdot \frac{4}{5} \cdot \frac{5}{10}}{\frac{3}{5} \cdot \frac{2}{5} \cdot \frac{4}{5} \cdot \frac{5}{10} + \frac{2}{5} \cdot \frac{2}{5} \cdot \frac{1}{5} \cdot \frac{5}{10}}$$

$$= \frac{6}{7} \approx 0.8571$$

$$P(-|A=1, B=1, C=1) = \frac{P(A=1, B=1, C=1|-)\,P(-)}{P(A=1, B=1, C=1|+)\,P(+) \quad P(A=1, B=1, C=1|-)\,P(-)}$$

$$= \frac{\frac{2}{5} \cdot \frac{2}{5} \cdot \frac{1}{5} \cdot \frac{5}{10}}{\frac{3}{5} \cdot \frac{2}{5} \cdot \frac{4}{5} \cdot \frac{5}{10} + \frac{2}{5} \cdot \frac{2}{5} \cdot \frac{1}{5} \cdot \frac{5}{10}}$$

$$= \frac{1}{7}$$

c) $P(A=1) = \frac{5}{10} = \frac{1}{2}$

$P(B=1) = \frac{4}{10} = \frac{2}{5}$

$P(A=1, B=1) = \frac{2}{10} = P(A=1) \cdot P(B=1)$ $\therefore$ $A=1$ & $B=1$ are independent events

d) $P(A=1) = \frac{1}{2}$

$P(B=0) = \frac{6}{10} = \frac{3}{5}$

$P(A=1, B=0) = \frac{3}{10} = P(A=1)\,P(B=0)$ $\therefore$ $A=1$ & $B=0$ are independent events

e) $P(A=1|+) = \frac{3}{5}$

$P(B=1|+) = \frac{2}{5}$

$P(A=1|+)\,P(B=1|+) = \frac{6}{25}$

$P(A=1, B=1|+) = \frac{1}{5} \neq \frac{6}{25}$

$\therefore$ $A=1$ & $B=1$ given $+$ are not conditionally independent events