CSE 422
Assignment 2

Kazi Abrab Hossain
ID: 21201496
Sec: 10

1 NO Ques Ans

a) $P(C \cap P | S) = \frac{P(C \cap P \cap S)}{P(S)}$

$= \frac{0.25}{0.25+0.02+0.15+0.03}$

$= 5/9$

$P(C|S) = \frac{P(C \cap S)}{P(S)}$ —①

$= \frac{0.25+0.02}{0.45}$

$= 0.6$

$P(P|S) = \frac{P(P \cap S)}{P(S)}$ —②

$= \frac{0.25+0.15}{0.45}$

$= 8/9$

① x ②

$0.6 \times 8/9 = 8/15$

$P(C \cap P|S) \neq P(C|S) * P(P|S)$

∴ Cheat and pass not conditionally independent given study.

b) $P(P \cup C) = P(P) + P(C) - P(P \cap C)$

$= (0.25+0.15+0.10+0.13) + (0.25+0.02 + 0.10+0.22) - (0.25+0.10)$

$= 0.87$

2 NO Ques Ans

a) $P(C') = (0.12+0.04+0.10+0.07)$

$= 0.33$

b) $P(C' \mid R' \cap Co') = \dfrac{P(C') \cap P(R') \cap P(Co')}{P(R' \cap Co')}$

$= \dfrac{0.07}{0.04+0.07}$

$= 7/11$

c) $P(R' \mid C') = \dfrac{P(C' \cap R')}{P(C')}$

$= \dfrac{0.03+0.07}{0.26+0.10+0.03+0.07}$

$= 5/23$

d) $P(R' \cup C) = P(R') + P(C) - P(R' \cap C)$

$= (0.06+0.04+0.03+0.07) + (0.32+0.12 + 0.06+0.04) - (0.06+0.04)$

$= 0.64$

3 NO Ques Ans

a) $P(fb \mid LH) = \frac{P(fb \cap LH)}{P(LH)}$

$= \frac{0.15}{0.24+0.15+0.15}$

$= 5/18$

b) $P(RH \mid C) = \frac{0.1}{0.24+0.1} = 5/17$

c) $P(fb \cap C) = P(fb) \times P(C)$

$= (0.15+0.1) \times (0.24+0.1)$

$= 0.085$

d) $P(RH \cup LH) = P(RH) + P(LH) - P(RH \cap LH)$

$= (0.1+0.1+0.26) + (0.24+0.15+0.15) - 0$

$= 1$

e) $P(fb \cap RH) = 0.1$

$= P(Fb) \times P(RH)$

$= (0.15+0.1) \times (0.1+0.1+0.26)$

$= 0.115$

$P(fb \cap RH) \neq P(fb) \times P(RH)$

Playing football depends on being right hand.

4 NO Ques Ans

a) $P(Cd|L) = 0.96$

$P(Cd|L') = 0.95$

$P(L) = 0.02$

$P(Cd'|L) = 0.04$

$P(Cd'|L') = 0.05$

$P(L') = 0.98$

$$P(L|Cd) = \frac{P(Cd|L)\,P(L)}{P(Cd)}$$

$$= \frac{0.0192}{P(Cd)}$$

$$P(L'|Cd) = \frac{P(Cd|L')\,P(L')}{P(Cd)}$$

$$= \frac{0.931}{P(Cd)}$$

$P(L'|Cd) > P(L|Cd)$

∴ The person did not lie.

b) $P(Cd) = P(Cd \cap L) \times P(Cd \cap L')$

$P(Cd \cap L') = P(Cd|L) \cdot P(L)$

$= 0.0192$

$$P(Cd \cap L') = P(Cd \mid L') \cdot P(L')$$
$$= 0.95 \times 0.98$$
$$= 0.931$$

$$P(Cd) = 0.192 + 0.931$$
$$= 0.9502$$

$$P(L' \mid Cd) = \frac{P(Cd \cap L')}{P(Cd)}$$
$$= 0.9798$$

5 NO Ques Ans

The issue with bayes theorm is that the features are conditionally dependent. If we want the probability of event happening, other erent need have to occure. If we want to find the probability of multiple feature given the outcome then calculating that probability is difficult. for example.

A	B	C	D
5	95	70	Yes
10	80	32	No
20	80	65	No
10	85	75	Yes

If we want to calculate $P(A \mid B \cap C \cap D)$ we need to know $P(A)$ but also $P(B \cap C \cap D \mid A)$. naive Bayes

Solves this by assuming all features are conditionally independent given outcome.

GNO Ques Ans

a)

$P(Y|S,M,N,T) = P(S,M,N,T|Y) * P(Y)$

$P(S|Y) = 2/5 \quad P(T|Y) = 3/5$

$P(M|Y) = 3/5 \quad P(Y) = 5/8$

$P(N|Y) = 4/5$

$P(Y|S,M,N,T) = 2/5 \times 3/5 \times 4/5 \times 3/5 \times 5/8$

$= 0.072$

$P(N_o|S,M,N,T) = P(S,M,N,T|N_o)\, P(N_o)$

$= 1 \times 2/3 \times 0 \times 1/3 \times 3/8$

$= 0$

∴ Player is going to play tennis.

b) $P(Y|OC,H,N,T) = P(OC,H,N,T|Y) \times P(Y)$

$= 2/5 \times 0 \times 4/5 \times 3/5 \times 5/8$

$= 0$

$P(N|OC,H,N,T) = P(OC,H,N,T|N) \times P(N)$

$= 0 \times 1/3 \times 0 \times 1/3 \times 3/8$

$= 0$

We cannot make prediction.

7 NO Ques Ans

a) Entropy (edible) $= -\frac{9}{16}\log_2\left(\frac{9}{16}\right) - \frac{7}{16}\log_2\left(\frac{7}{16}\right)$

$= 0.9869$

b] Information (color) $= 0.989 - \left(\frac{13}{16} \times \left(-\frac{8}{13}\log_2(8/13) - \frac{5}{13}\log_2(5/13)\right) + \frac{3}{16} \times \left(-\frac{1}{3}\log_2\left(\frac{1}{3}\right) - \frac{2}{3}\log_2 \frac{2}{3}\right)\right)$

$= 0.03$

Information (Size) $= 0.989 - \left(\frac{8}{16} \times \left(-\frac{6}{8}\log_2 \frac{6}{8} - \frac{2}{8}\log_2\left(\frac{2}{8}\right)\right)\right.$

$\left. + \frac{8}{16} \times \left(-\frac{3}{8}\log_2 \frac{3}{8} - \frac{5}{8}\log_2 \frac{5}{8}\right)\right)$

$= 0.11$

Information (Shape) $= 0.989 - \left(\frac{12}{16} \times \left(-\frac{6}{12}\log_2\left(\frac{6}{12}\right) - \frac{6}{12}\log_2\left(\frac{6}{12}\right)\right) + \frac{4}{16}\left(-\frac{3}{4}\log_2(3/4) - \frac{1}{4}\log_2\frac{1}{4}\right)\right)$

$= 0.04$

∴ Size is better feature as it has ~~more~~ most information gain.