Nirav

①

Q.5 We can write Naive Bayes assumption by:

$P(A,B,C,P) = P(P) \cdot P(A|P) \cdot P(B|P) \cdot P(C|P)$

A	B	C	P	Tweets	Probability
f	f	f	f	10	0.05
f	f	f	t	18	0.09
f	f	t	f	40	0.2
f	f	t	t	69	0.34
f	t	f	f	1	0.005
f	t	f	t	2	0.01
f	t	t	f	2	0.01
f	t	t	t	5	0.025
t	f	f	f	5	0.025
t	f	f	t	6	0.03
t	f	t	f	18	0.09
t	f	t	t	19	0.095
t	t	t	f	2	0.01
t	t	t	t	3	0.015
			total	200	1

for P(P)

P	P(P)	
t	18 + 69 + 2 + 5 + 6 + 19 + 3 /200	= 0.61
f	10 + 40 + 1 + 2 + 5 + 18 + 2 /200	= 0.39

for P(A|P)

A	P	P(A|P)	
t	t	6 + 19 + 3 / 122	= 0.23
t	f	5 + 18 + 2 / 78	= 0.32
f	t	18 + 69 + 2 + 5 / 122	= 0.77
f	f	10 + 40 + 2 + 1 / 78	= 0.68

for P(B|P)

B	P	P(B|P)	
t	t	2+5+3 / 122	= 0.081
t	f	1+2+2 / 78	= 0.064
f	t	18+69 + 6+ 19 / 122	= 0.92
f	f	10+ 40+ 5+18 / 78	= 0.94

for P(C|P)

C	P	P(C|P)	
t	t	69 + 5 + 19 + 3 / 122	= 0.79
t	f	40 + 2 + 18 + 2 / 78	= 0.79
f	t	18 + 2 + 6 / 122	= 0.21
f	f	10 + 1 + 5 / 78	= 0.20

B) conditional probability measure probability of an Event based on another occured Event.

here CP $P(P=t \mid A=f, B=t, C=f)$ means word A and C are not there in comment and B presents in comment.

$P(P=t \mid A=f, B=t, C=f) =$

$$P(P=t, A=f, B=t, C=f) / P(A=f, B=t, C=f)$$

using Naive bayes assumption

$= P(P=t, A=f, B=t, C=f) =$

$= P(P=t)\, P(A=f \mid P=t)\, P(B=t \mid P=t)\, P(C=f \mid P=t)$
$= \underline{0.00798}$

$P(A=f, B=t, C=f) = P(P=t, A=f, B=t, C=f)$
$+$
$P(P=f, A=f, B=t, C=f)$

$= P(P=t, A=f, B=t, C=f) + P(P=f)\, P(A=f \mid P=f)$
$P(B=t \mid P=f)\; P(C=f \mid P=f)$

$= 0.00798 + 0.00339$

$= 0.01137$

$\therefore P(P=t \mid A=f, B=t, C=f) = \underline{0.702}$

c) most likely value for P

$\arg\max P(P=x \mid A=f, B=t, C=f) =$

$\arg\max \left(P(P=x, A=f, B=t, C=f) / P(A=f, B=t, C=f) \right)$

$= \arg\max P(P=x)\, P(A=f \mid P=x)\, P(B=t \mid P=x)\, P(C=f \mid P=x)$ — ①

① is Naive bayes assumption.

Now assume ① as $z(x)$

for $x=t \Rightarrow z(t) = P(P=t)\, (A=f \mid P=t)\, P(B=t \mid P=t)\, P(C=f \mid P=t)$

$= \underline{0.00798}$

for $x=f \Rightarrow z(f) = P(P=f)\, P(A=f \mid P=f)\, P(B=t \mid P=f)\, P(C=f \mid P=f)$

$= \underline{0.00339}$

here $z(t) > z(f)$

So most likely value here for P is $\underline{t}$.

d) CP of Joint distribution model

$P(P=t \mid A=f, B=t, C=f) =$

$P(P=t, A=f, B=t, C=f) / P(A=f, B=t, C=f)$

$= P(P=t, A=f, B=t, C=f) /$
$(P(P=t, A=f, B=t, C=f) +$
$P(P=f, A=f, B=t, C=f))$

$= \underline{0.667}$

e) CP of fully independent model

$P(P=t \mid A=f, B=t, C=f) =$

$P(P=t, A=f, B=t, C=f) / P(A=f, B=t, C=f)$

$= P(P=t) P(A=f) P(B=t) P(C=f) /$

$P(A=f) P(B=t) P(C=f)$

$= \underline{0.61}$