Consider the following dataset of a credit card promotion database. The credit card company has authorized a new life insurance promotion similar to the existing one. We are interested in building a classification data mining model for deciding whether to send the customer promotional material.

Customer ID	**Magazine Promotion**	**Watch Promotion**	**Credit Card Insurance**	**Sex**	**Life Insurance Promotion**
1	Y	N	N	M	N
2	Y	Y	Y	F	Y
3	N	N	N	M	N
4	Y	Y	Y	M	Y
5	Y	N	N	F	Y
6	N	N	N	F	N
7	Y	Y	Y	M	Y
8	N	N	N	M	N
9	Y	Y	Y	M	N
10	N	Y	N	F	Y

(a) Build a Naive Bayes classifier for this dataset, by filling in the following with counts and probabilities.

		Life Insurance Promotion (*LIP*)	
		Y	N
Magazine Promotion (*MP*)	Y	4	2
	N	1	3
***P*(*MP* = ? | *LIP* = ?)**	Y	4/5	2/5
	N	1/5	3/5

		Life Insurance Promotion (*LIP*)	
		Y	N
Watch Promotion (*WP*)	Y	4	1
	N	1	4
***P*(*WP* = ? | *LIP* = ?)**	Y	4/5	1/5
	N	1/5	4/5

		Life Insurance Promotion (*LIP*)	
		Y	N
Credit Card Insurance (*CCI*)	Y	3	1
	N	2	4
***P*(*CCI* = ? | *LIP* = ?)**	Y	3/5	1/5
	N	2/5	4/5

		Life Insurance Promotion (*LIP*)	
		Y	N
Sex (*S*)	M	2	4
	F	3	1
***P*(*S* = ? | *LIP* = ?)**	M	2/5	4/5
	F	3/5	1/5

	Life Insurance Promotion (*LIP*)	
	Y	N
Counts	5	5
***P*(*LIP* = ?)**	5/10	5/10

(b) Let X = (MP = Y, WP = Y, CCI = N, S = F). Calculate the conditional probabilities $P(LIP = \text{Y} \mid X)$ and $P(LIP = \text{N} \mid X)$ in terms of $P(X)$.

Hint: Use Bayes theorem

$$P(H \mid X) = \frac{P(X \mid H)P(H)}{P(X)}$$

and assume class conditional independence, i.e.,

$$P(X \mid H) = P(MP = \text{Y} \mid H) \times P(WP = \text{Y} \mid H) \times P(CCI = \text{N} \mid H) \times P(S = \text{F} \mid H)$$

where H is LIP = Y or LIP = N.

Solution:

Let $H_1 = (LIP = \text{Y})$,

$$P(X \mid H_1) = P(MP = \text{Y} \mid H_1) \times P(WP = \text{Y} \mid H_1) \times P(CCI = \text{N} \mid H_1) \times P(S = \text{F} \mid H_1)$$

$$= \frac{4}{5} \times \frac{4}{5} \times \frac{2}{5} \times \frac{3}{5} = \frac{96}{625}$$

$$P(LIP = \text{Y} \mid X) = P(H_1 \mid X)$$

$$= \frac{P(X \mid H_1)P(H_1)}{P(X)}$$

$$= \frac{\frac{96}{625} \times \frac{5}{10}}{P(X)} = \frac{\frac{48}{625}}{P(X)} = \frac{0.0768}{P(X)}$$

Let $H_2 = (LIP = \text{N})$,

$$P(X \mid H_2) = P(MP = \text{Y} \mid H_2) \times P(WP = \text{Y} \mid H_2) \times P(CCI = \text{N} \mid H_2) \times P(S = \text{F} \mid H_2)$$

$$= \frac{2}{5} \times \frac{1}{5} \times \frac{4}{5} \times \frac{1}{5} = \frac{8}{625}$$

$$P(LIP = \text{N} \mid X) = P(H_2 \mid X)$$

$$= \frac{P(X \mid H_2)P(H_2)}{P(X)}$$

$$= \frac{\frac{8}{625} \times \frac{5}{10}}{P(X)} = \frac{\frac{4}{625}}{P(X)} = \frac{0.0064}{P(X)}$$

(c) Use the Naive Bayes classifier obtained in part (a) and results in part (b) to determine the value of Life Insurance Promotion for the following instance:

Magazine Promotion = Y
Watch Promotion = Y
Credit Card Insurance = N
Sex = F
Life Insurance Promotion = ?

Explain your answers clearly.

Solution:

Since $P(LIP = \text{Y} \mid X) > P(LIP = \text{N} \mid X)$, therefore we predict that LIP = Y.