TASK 1

i) In the given data :

$n_+ = 4, \quad n_- = 4$

$\forall i, \quad s_i = 2$

The required values are in the following table :

i	j	k	n_{ijk}	n_k	s_i	$P(A_i = v_j \mid C_k) = \frac{n_{ijk}+1}{n_k + s_i}$
1	0	−	2	4	2	0.5
1	0	+	1	4	2	0.333
1	1	−	2	4	2	0.5
1	1	+	3	4	2	0.666
2	0	−	3	4	2	0.666
2	0	+	1	4	2	0.333
2	1	−	1	4	2	0.333
2	1	+	3	4	2	0.666
3	0	−	2	4	2	0.5
3	0	+	2	4	2	0.5
3	1	−	2	4	2	0.5
3	1	+	2	4	2	0.5
4	0	−	4	4	2	0.833
4	0	+	1	4	2	0.333
4	1	−	0	4	2	0.1666
4	1	+	3	4	2	0.666

ii.)

	A_1	A_2	A_3	A_4	C
x	1	1	0	0	

A naive Bayes classifier predicts C as follows

$$\boxed{\underset{k}{\operatorname{Argmax}}\ P(C_k \mid A_1 = 1, A_2 = 1, A_3 = 0, A_4 = 0)}$$

$$P(C_+ \mid A_1=1, A_2=1, A_3=0, A_4=0) =$$

$$\frac{P(A_1=1, A_2=1, A_3=0, A_4=0 \mid C_+) \times P(C_+)}{P(A_1=1, A_2=1, A_3=0, A_4=0)}$$

(By Baye's Thm)

Since the denominator is a normalizing factor, we can ignore it.

Also, The attributes are assumed to be independent

Hence,

$$= P(A_1=1 \mid C_+) \times P(A_2=1 \mid C_+) \times P(A_3=0 \mid C_+) \times P(A_4=0 \mid C_+) \times P(C_+)$$

By using Laplacian estimates from Task 1:

$$= \frac{2}{3} \times \frac{2}{3} \times \frac{1}{2} \times \frac{1}{3} \times \frac{1}{2} = \frac{1}{27} = 0.037$$

Similarly,

$$P(C_- \mid A_1 = 1, A_2 = 1, A_3 = 0, A_4 = 0)$$

$$= P(A_1 = 1 \mid C_-) \times P(A_2 = 1 \mid C_-) \times P(A_3 = 0 \mid C_-) \times P(A_4 = 0 \mid C_-) \times P(C_-)$$

$$= 1/2 \times 1/3 \times 1/2 \times 5/6 \times 1/2 = 5/144 = 0.034$$

Since $P(C_+ \mid A_1=1, A_2=1, A_3=0, A_4=0) > P(C_- \mid A_1=1, A_2=1, A_3=0, A_4=0)$

$\boxed{\text{Predicted class} = +}$