

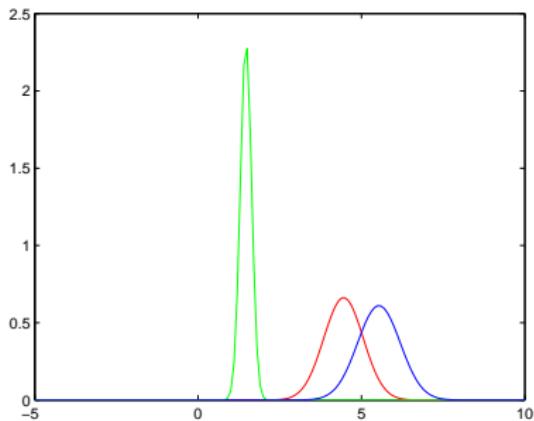
Ans(1-1)

$P(C_0) = 0.39$, $P(C_1) = 0.33$ and $P(C_2) = 0.28$ denote the probability of the cluster 0, cluster 1 and cluster 2 respectively. X is the value of petal length and those conditional probabilities generated by EM algorithm are described as follows:

$$P\{X|C_0\} = \frac{1}{\sqrt{2\pi}\sigma_0} \exp^{-\frac{(x-\mu_0)^2}{2\sigma_0^2}}, \text{ } \mu_0 \text{ is } 4.4513 \text{ and } \sigma_0 \text{ is } 0.6016 \text{ (Red).}$$

$$P\{X|C_1\} = \frac{1}{\sqrt{2\pi}\sigma_1} \exp^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}, \text{ } \mu_1 \text{ is } 1.464 \text{ and } \sigma_1 \text{ is } 0.1717 \text{ (Green).}$$

$$P\{X|C_2\} = \frac{1}{\sqrt{2\pi}\sigma_2} \exp^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}}, \text{ } \mu_2 \text{ is } 5.5353 \text{ and } \sigma_2 \text{ is } 0.6527 \text{ (Blue).}$$



Ans(1-1)

We want to know:

$$P\{C_0|X = 2.0\}, P\{C_1|X = 2.0\}, P\{C_2|X = 2.0\}.$$

$$P\{C_0|X = 5.0\}, P\{C_1|X = 5.0\}, P\{C_2|X = 5.0\}.$$

$$P\{C_0|X = 6.0\}, P\{C_1|X = 6.0\}, P\{C_2|X = 6.0\}.$$

Using Bayes' rule,

$$P\{C_0|X = 2.0\} = \frac{P\{X = 2.0|C_0\} \cdot P\{C_0\}}{P\{X = 2.0\}}$$

$$\begin{aligned} P\{X = 2.0\} &= P\{X = 2.0|C_0\} \cdot P\{C_0\} + P\{X = 2.0|C_1\} \cdot P\{C_1\} + P\{X = 2.0|C_2\} \cdot P\{C_2\} \\ &= 1.70276 \times 10^{-6} \times 0.39 + 1.62513 \times 10^{-4} \times 0.33 + 2.71478 \times 10^{-9} \times 0.28 \\ &= 5.429 \times 10^{-5} \end{aligned}$$

$$P\{C_0|X = 2.0\} = \frac{1.70276 \times 10^{-6} \times 0.39}{5.429 \times 10^{-5}} \simeq 1.22\%$$

$$P\{C_1|X = 2.0\} = \frac{1.62513 \times 10^{-4} \times 0.33}{5.429 \times 10^{-5}} \simeq 98.78\%$$

$$P\{C_2|X = 2.0\} = \frac{2.71478 \times 10^{-9} \times 0.28}{5.429 \times 10^{-5}} \simeq 0\%$$

So, $X = 2.0$ belongs to the cluster 1 with the highest probability.

Ans(1-1)

$$P\{C_0|X = 5.0\} = \frac{P\{X = 5.0|C_0\} \cdot P\{C_0\}}{P\{X = 5.0\}}$$

$$\begin{aligned} P\{X = 5.0\} &= P\{X = 5.0|C_0\} \cdot P\{C_0\} + P\{X = 5.0|C_1\} \cdot P\{C_1\} + P\{X = 5.0|C_2\} \cdot P\{C_2\} \\ &= 4.34167 \times 10^{-3} \times 0.39 + 0.0 \times 0.33 + 4.39394 \times 10^{-3} \times 0.28 \\ &= 2.9235545 \times 10^{-3} \end{aligned}$$

$$P\{C_0|X = 5.0\} = \frac{4.34167 \times 10^{-3} \times 0.39}{2.9235545 \times 10^{-3}} \simeq 57.91\%$$

$$P\{C_1|X = 5.0\} = \frac{0 \times 0.33}{2.9235545 \times 10^{-3}} \simeq 0.00\%$$

$$P\{C_2|X = 5.0\} = \frac{4.39394 \times 10^{-3} \times 0.28}{2.9235545 \times 10^{-3}} \simeq 42.09\%$$

$$P\{C_0|X = 6.0\} = \frac{P\{X = 6.0|C_0\} \cdot P\{C_0\}}{P\{X = 6.0\}}$$

$$\begin{aligned} P\{X = 6.0\} &= P\{X = 6.0|C_0\} \cdot P\{C_0\} + P\{X = 6.0|C_1\} \cdot P\{C_1\} + P\{X = 6.0|C_2\} \cdot P\{C_2\} \\ &= 2.36201 \times 10^{-4} \times 0.39 + 0.0 \times 0.33 + 4.71783 \times 10^{-3} \times 0.28 \\ &= 1.413111079 \times 10^{-3} \end{aligned}$$

$$P\{C_0|X = 6.0\} = \frac{2.36201 \times 10^{-4} \times 0.39}{1.413111079 \times 10^{-3}} \simeq 6.52\%$$

$$P\{C_1|X = 6.0\} = \frac{0 \times 0.33}{1.413111079 \times 10^{-3}} \simeq 0.00\%$$

$$P\{C_2|X = 6.0\} = \frac{4.71783 \times 10^{-3} \times 0.28}{1.413111079 \times 10^{-3}} \simeq 93.48\%$$

Ans(1-2)

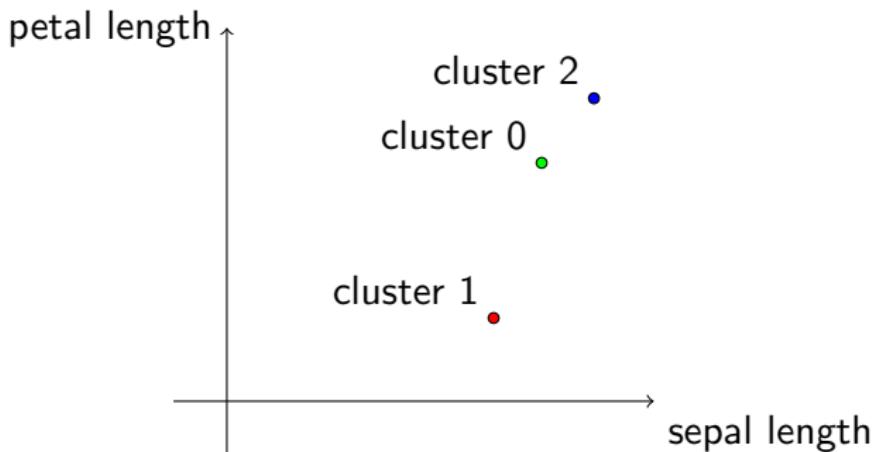
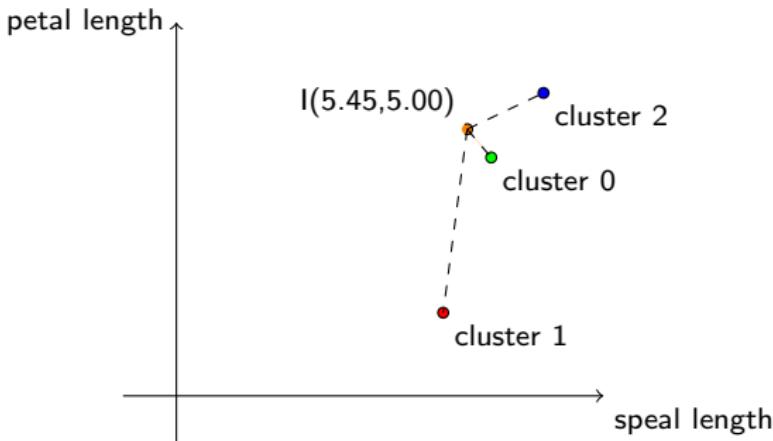


Table: Simple K Means Results

Attribute	Cluster 0	Cluster 1	Cluster 2
sepal length	5.9086	5.0057	6.8846
petal length	4.4759	1.5623	5.6769

Ans(1-2)



$$\|(I, \text{cluster}0)\| = \sqrt{(5.45 - 5.9086)^2 + (5.00 - 4.4759)^2} = 0.6943$$

$$\|(I, \text{cluster}1)\| = \sqrt{(5.45 - 5.0057)^2 + (5.00 - 1.5623)^2} = 3.468$$

$$\|(I, \text{cluster}2)\| = \sqrt{(5.45 - 6.8846)^2 + (5.00 - 5.6769)^2} = 1.5834$$

Since $\|(I, \text{cluster}0)\|$ has the smallest distance from cluster 0, this instance is clustered to the cluster 0.