Chapter 1

1.1 silhouette coefficient

To calculate the silhouette coefficient of a single example in our dataset, we can apply the following three steps:

1. Calculate the cluster cohesion, $a^{(i)}$, as the average distance between an example, $\boldsymbol{x}^{(i)}$, and all other points in the same cluster.

2. Calculate the cluster separation, $b^(i)$, from the next closest cluster as the average distance between the example, $\boldsymbol{x}^{(i)}$, and all examples in the nearest cluster.

3. Calculate the silhouette, $s^{(i)}$, as the difference between cluster cohesion and separation divided by the greater of the two, as shown here:

$$s^{(i)} = \frac{b^{(i)} - a^{(i)}}{\max\{b^{(i)}, a^{(i)}\}} \tag{1.1}$$