

Initial Cluster Centres were randomly generated. For some sequences, it was observed that the classification accuracy changed drastically but for most, the accuracy stayed at around 65%. One such initial set of cluster centres is

[[6.1 2.8 4. 1.3]

[7.7 3. 6.1 2.3]

[6.3 3.4 5.6 2.4]]

This produces output with jaccard index = 0.6958587915818059, which means jaccard distance between test set and prediction set = 0.3041412084181941

Whereas

[[4.4 2.9 1.4 0.2]

[5.7 4.4 1.5 0.4]

[5.5 2.5 4. 1.3]]

Jaccard Index = 0.4768, which means jaccard distance = 0.5232 between the training and prediction set.