Dataset: Indian and Srilankan Nationality

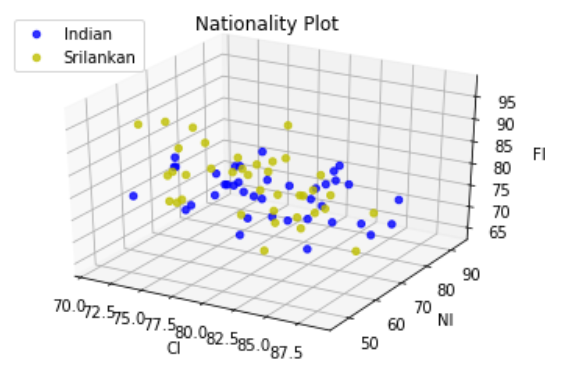
Number of Samples = 77

Figure: Unclustered Data

Test Dataset Size = 25%

Graph of model inertia vs number of model clusters.

From this graph using elbow method this is evident that classifying Nationality from given dataset is highly possible by clustering as the optimum number of clusters is 2. For cluster size of 2 the inertia is 73080.33896565421

which can be reduced increasing sample size.

[omitted the elbow method entirely]

Attributes used(6): Age, Nationality, Sex, HI, FI, NI

Cross-tabulated output of the dataset on given labels of nationality (srilankan & Indian) and the predicted output by using cluster centroids.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Predicted Class | |
|  |  | Indian | Srilankan |
| Actual Class | Indian | 9 | 3 |
| Srilankan | 3 | 5 |

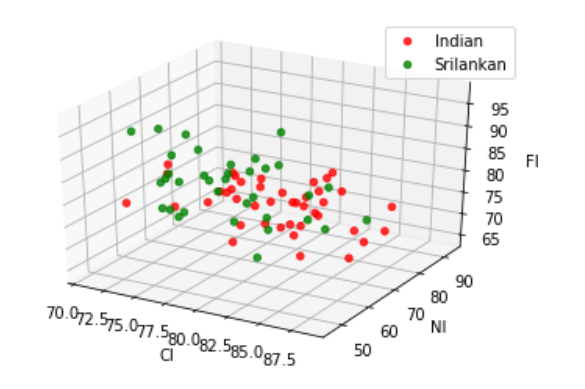


Fig. Data points after clustering

Using the cross tabulation results, we predicted the red points as Indian class and the green points as Srilankan class.

Accuracy = 70%  
Precision for predicting Indian =75%  
Precision for predicting Srilankan =62.5%  
Recall for predicting Indian = 75%  
Recall for predicting Female = 62.5%

Output Data

PIN Age Sex CI NI FI prediction

0 5 18 2 77.876106 64.285714 84.191176 Srilankan

1 13 20 2 79.056047 55.454545 83.870968 Srilankan

2 23 20 2 75.561798 53.781513 86.524823 Srilankan

3 44 20 2 83.582090 67.391304 81.224490 Srilankan

4 45 20 2 77.989130 70.329670 75.378788 Srilankan

5 94 20 2 87.134503 56.701031 79.467681 Srilankan

6 96 18 2 82.005141 82.417582 76.872964 Srilankan

7 97 22 2 72.533333 69.411765 69.550173 Srilankan

8 98 21 2 76.881720 78.494624 82.142857 Srilankan

9 120 20 2 71.478261 55.952381 77.042802 Indian

10 147 18 2 88.650307 63.888889 74.809160 Indian

11 148 20 2 79.706345 75.641026 76.811594 Indian

12 149 18 2 82.954545 67.435897 73.929961 Indian

13 150 19 2 80.280112 81.159420 72.075472 Indian

14 151 19 2 77.472527 77.631579 76.258993 Indian

15 152 21 2 84.567901 80.303030 68.973384 Indian

16 153 19 2 78.116343 72.727273 73.977695 Indian

17 174 19 2 86.705202 81.818182 69.536424 Indian

18 175 20 2 86.472303 85.507246 73.703704 Indian

19 176 20 2 81.936416 72.222222 64.569776 Indian

20 54 20 1 72.357724 72.093023 69.565217 Srilankan

21 95 28 1 75.456919 68.041237 80.000000 Srilankan

22 129 23 1 76.519337 71.618037 65.034483 Indian

23 145 20 1 76.377953 71.951220 77.142857 Indian

24 190 37 1 83.943662 74.418605 84.030418 Indian

25 198 33 1 72.323760 66.666667 80.377358 Indian

26 201 19 1 79.063361 60.674157 86.131387 Indian

27 215 20 1 78.092784 82.500000 65.306122 Indian

28 218 36 1 79.729730 69.148936 72.013652 Indian

29 219 19 1 76.487252 63.043478 77.407407 Indian

30 222 40 1 75.000000 73.913043 75.170068 Indian

31 223 38 1 83.149171 82.022472 76.712329 Indian

32 226 19 1 85.635359 60.000000 81.071429 Indian

33 227 39 1 77.486911 68.888889 76.632302 Indian

34 229 39 1 70.810811 73.333333 79.467681 Indian

35 235 42 1 76.633166 69.135802 77.543860 Indian

36 237 39 1 77.777778 68.965517 70.671378 Indian

37 240 42 1 81.198910 85.542169 77.528090 Indian

38 241 38 1 78.933333 93.333333 70.508475 Indian

39 8 19 1 81.492537 66.037736 87.272727 Srilankan

40 30 20 1 71.542553 67.368421 89.887640 Srilankan

41 35 19 1 77.747989 78.350515 80.398671 Srilankan

42 39 19 1 78.042328 62.608696 83.576642 Srilankan

43 41 18 1 78.249337 70.297030 82.746479 Srilankan

44 47 19 1 77.685950 74.074074 75.265018 Srilankan

45 65 21 1 82.708934 71.264368 73.776224 Srilankan

46 67 21 1 75.000000 66.326531 87.272727 Srilankan

47 70 21 1 81.163435 63.736264 73.426573 Srilankan

48 72 21 1 71.067416 69.512195 76.727273 Srilankan

49 73 19 1 75.197889 67.391304 81.227437 Srilankan

50 76 19 1 72.679045 63.218391 73.897059 Srilankan

51 77 19 1 75.623269 59.615385 93.172691 Srilankan

52 142 21 1 79.862941 86.835443 75.444840 Indian

53 143 18 1 82.271468 79.250000 74.733096 Indian

54 212 19 1 76.546392 74.698795 77.978339 Indian

55 255 22 1 80.053908 66.304348 83.453237 Indian

56 1 20 2 77.747253 65.789474 85.555556 Srilankan

57 9 21 2 71.934605 70.000000 83.269962 Srilankan

58 40 18 2 72.527473 54.330709 94.052045 Srilankan

59 42 19 2 83.013699 59.813084 97.233202 Srilankan

60 43 28 2 73.463687 65.934066 79.389313 Srilankan

61 66 21 2 84.112150 46.511628 75.289575 Srilankan

62 68 22 2 86.764706 74.193548 74.910394 Srilankan

63 69 20 2 72.752044 64.285714 80.303030 Srilankan

64 71 20 2 81.142857 63.333333 76.277372 Srilankan

65 74 21 2 77.142857 69.411765 70.955882 Srilankan

66 75 18 2 74.927114 56.470588 77.235772 Srilankan

67 136 23 2 81.137725 73.428571 68.525896 Indian

68 137 19 2 73.927577 62.352941 75.272727 Indian

69 138 20 2 85.714286 72.857143 66.023166 Indian

70 139 21 2 76.553672 85.714286 69.455104 Indian

71 140 19 2 82.005900 67.924528 73.584906 Indian

72 141 20 2 80.691643 76.285714 74.803150 Indian

73 144 20 2 84.219653 67.948718 76.798144 Indian

74 177 20 2 81.104651 73.076923 76.284585 Indian

75 178 23 2 80.392157 81.428571 74.131274 Indian

76 179 21 2 73.333333 86.776860 73.484848 Indian