Purity-based evaluation for K-means

In this analysis, let **k = r = 3**. The contingency table is shown as follow:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| NO. |  | T1  iris-setosa | T2iris-versicolor | T3 iris-virginica | n | Purity Score |
| 1 | C1 | 50 | 0 | 0 | 50 | 0.887 |
| C2 | 0 | 47 | 14 | 47 |
| C3 | 0 | 3 | 36 | 36 |
|  |  |  |  | 133 |
| 2 | C1 | 0 | 2 | 36 | 36 | 0.893 |
| C2 | 50 | 0 | 0 | 50 |
| C3 | 0 | 48 | 14 | 48 |
|  |  |  |  | 134 |
| 3 | C1 | 0 | 48 | 14 | 48 | 0.893 |
| C2 | 50 | 0 | 0 | 50 |
| C3 | 0 | 2 | 36 | 36 |
|  |  |  |  | 134 |
| 4 | C1 | 50 | 0 | 0 | 50 | 0.893 |
| C2 | 0 | 2 | 36 | 36 |
| C3 | 0 | 48 | 14 | 48 |
|  |  |  |  | 134 |
| 5 | C1 | 0 | 3 | 36 | 36 | 0.887 |
| C2 | 0 | 47 | 14 | 47 |
| C3 | 50 | 0 | 0 | 50 |
|  |  |  |  | 133 |
| 6 | C1 | 0 | 48 | 14 | 48 | 0.893 |
| C2 | 50 | 0 | 0 | 50 |
| C3 | 0 | 2 | 36 | 36 |
|  |  |  |  | 134 |
| 7 | C1 | 50 | 0 | 0 | 50 | 0.893 |
| C2 | 0 | 2 | 36 | 36 |
| C3 | 0 | 48 | 14 | 48 |
|  |  |  |  | 134 |
| 8 | C1 | 0 | 3 | 36 | 36 | 0.887 |
| C2 | 0 | 47 | 14 | 47 |
| C3 | 50 | 0 | 0 | 50 |
|  |  |  |  | 133 |
| 9 | C1 | 0 | 47 | 50 | 50 | 0.667 |
| C2 | 26 | 3 | 0 | 26 |
| C3 | 24 | 0 | 0 | 24 |
|  |  |  |  | 100 |
| 10 | C1 | 0 | 48 | 14 | 48 | 0.893 |
| C2 | 0 | 2 | 36 | 36 |
| C3 | 50 | 0 | 0 | 50 |
|  |  |  |  | 134 |

The program is run for 10 times, and the best purity score is 0.893.