Multi-objective GenClust++ vs MCPSO-based algorithm

vs Elbow method for K-Means clustering

Table 1. Mean and standard deviation of ARI (higher the better) measured on the outputs of MCPSO, MGenClust++ and K-Means (over 30 independent runs).

|  |  |  |  |
| --- | --- | --- | --- |
| Dataset | MCPSO | MGenClust++ | K-Means |
| Glass | 0.4611 ± 0.0259 | 0.5309 ± 0.1856 | 0.7706 ± 0.0177 |
| Wdbc | 0.6291 ± 0.0666 | 0.6984 ± 0.0392 | 0.713 ± 0.0112 |
| Flame | 0.4251 ± 0.0577 | 0.4316 ± 0.052 | 0.4238 ± 0.0065 |
| Compound | 0.5404 ± 0.1482 | 0.7245 ± 0.0178 | 0.688 ± 0.103 |
| Pathbased | 0.4429 ± 0.0285 | 0.4486 ± 0.0169 | 0.4553 ± 0.0031 |
| Jain | 0.3614 ± 0.1172 | 0.5437 ± 0.0489 | 0.2803 ± 0.0 |
| S1 | 0.8952 ± 0.2446 | 0.9588 ± 0.0316 | 0.7646 ± 0.0471 |
| S3 | 0.5208 ± 0.0962 | 0.6472 ± 0.0315 | 0.6095 ± 0.0315 |
| DIM064 | 0.9682 ± 0.1563 | 0.8827 ± 0.2126 | 0.875 ± 0.0692 |
| DIM256 | 0.9999 ± 7.0E-4 | 0.8273 ± 0.3035 | 0.6225 ± 0.0789 |

Table 2. Mean and standard deviation of DB Index (lower the better) measured on the outputs of MCPSO, MGenClust++ and K-Means (over 30 independent runs).

|  |  |  |  |
| --- | --- | --- | --- |
| Dataset | MCPSO | MGenClust++ | K-Means |
| Glass | 0.5258 ± 0.0046 | 0.5095 ± 0.0033 | 0.5121 ± 1.0E-4 |
| Wdbc | 1.4731 ± 0.2437 | 1.1477 ± 0.1063 | 1.1087 ± 0.0033 |
| Flame | 0.9274 ± 0.2459 | 0.8781 ± 0.1334 | 0.7047 ± 0.0069 |
| Compound | 1.0167 ± 0.359 | 0.5893 ± 0.0637 | 0.6004 ± 0.1682 |
| Pathbased | 0.6895 ± 0.0442 | 0.7474 ± 0.0537 | 0.7135 ± 0.0111 |
| Jain | 0.6867 ± 0.0492 | 0.7818 ± 0.0265 | 0.6388 ± 0.0 |
| S1 | 0.5567 ± 0.4139 | 0.4627 ± 0.0755 | 0.5807 ± 0.1086 |
| S3 | 0.7208 ± 0.06 | 0.7854 ± 0.0498 | 0.7429 ± 0.052 |
| DIM064 | 0.213 ± 0.6173 | 0.3988 ± 0.3197 | 1.1892 ± 0.3649 |
| DIM256 | 0.0354 ± 0.0552 | 0.5676 ± 0.6385 | 1.328 ± 0.3549 |

Table 3. Mean and standard deviation of Silhouette coefficient (higher the better) measured on the outputs of MCPSO, MGenClust++ and K-Means (over 30 independent runs).

|  |  |  |  |
| --- | --- | --- | --- |
| Dataset | MCPSO | MGenClust++ | K-Means |
| Glass | 0.5216 ± 0.0048 | 0.5768 ± 0.0241 | 0.5793 ± 3.0E-4 |
| Wdbc | 0.2999 ± 0.0807 | 0.385 ± 0.0179 | 0.3907 ± 3.0E-4 |
| Flame | 0.3676 ± 0.1083 | 0.3963 ± 0.0148 | 0.4411 ± 0.0014 |
| Compound | 0.3585 ± 0.1302 | 0.5912 ± 0.0249 | 0.5849 ± 0.0715 |
| Pathbased | 0.5292 ± 0.0234 | 0.4936 ± 0.045 | 0.5351 ± 2.0E-4 |
| Jain | 0.5028 ± 0.0023 | 0.5078 ± 8.0E-4 | 0.5035 ± 0.0 |
| S1 | 0.6346 ± 0.1812 | 0.6804 ± 0.0232 | 0.5925 ± 0.0396 |
| S3 | 0.4519 ± 0.0246 | 0.4461 ± 0.0173 | 0.4495 ± 0.0198 |
| DIM064 | 0.9253 ± 0.1748 | 0.8717 ± 0.1443 | 0.751 ± 0.08 |
| DIM256 | 0.9812 ± 0.0084 | 0.8439 ± 0.2122 | 0.5967 ± 0.0751 |

Table 4. Mean and standard deviation of the average number of clusters (over 30 independent runs) for MCPSO, MGenClust++ and K-Means

|  |  |  |  |
| --- | --- | --- | --- |
| Dataset | MCPSO | MGenClust++ | K-Means |
| Glass | 6.7667 ± 0.423 | 3.3 ± 0.781 | 3.0 ± 0.0 |
| Wdbc | 2.8 ± 0.4 | 2.1 ± 0.3 | 2.0 ± 0.0 |
| Flame | 2.9 ± 0.5972 | 2.7333 ± 0.4422 | 4.0 ± 0.0 |
| Compound | 7.3333 ± 4.1899 | 3.2667 ± 0.5121 | 3.0 ± 0.0 |
| Pathbased | 2.7333 ± 0.4422 | 3.5333 ± 0.5617 | 3.0 ± 0.0 |
| Jain | 2.8 ± 0.4 | 2.0333 ± 0.1795 | 3.0 ± 0.0 |
| S1 | 14.4667 ± 3.3539 | 15.7 ± 0.9713 | 12.0 ± 0.0 |
| S3 | 9.6 ± 2.2301 | 19.2667 ± 2.0645 | 13.0 ± 0.0 |
| DIM064 | 15.8333 ± 1.5074 | 15.1667 ± 2.5701 | 17.3667 ± 1.2243 |
| DIM256 | 16.0333 ± 0.1795 | 15.1 ± 2.8792 | 10.8 ± 0.8327 |
| Average difference between the real and the detected numbers of clusters | 8.1267 ± 1.3723 | 1.5367 ± 1.0168 | 7.1167 ± 0.2057 |

Fig. 1. Comparative average results between MCPSO, MGenClust++ and K-Means on 10 datasets based on ARI, DB Index and Silhouette Coefficient