

Table B4. Performance (%) evaluation of different datasets based on the NMI metric. We have highlighted the values of the best-performing method in **bold**, and the second-best method is marked with an underline. • indicates whether proposed method is statistically superior to the compared methods according to the pairwise t-test at 0.05 significance level.

Method	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	Average
CEAM (TKDE'24)	•5.6 \pm 10	•36.2 \pm 26	•16.7 \pm 4	•27.4 \pm 1	•60.1 \pm 10	•4.3 \pm 3	•18.0 \pm 2	•19.0 \pm 5	•14.3 \pm 4	•8.8 \pm 5	21.0 \pm 8
CES ² L (AIJ'19)	•3.4 \pm 5	•9.3 \pm 10	•19.0 \pm 4	•27.9 \pm 2	•45.1 \pm 14	•12.3 \pm 5	•12.0 \pm 2	•15.2 \pm 7	<u>15.7</u> \pm 3	•10.2 \pm 6	17.0 \pm 6
CES ² Q (AIJ'19)	•2.5 \pm 4	•11.5 \pm 8	•17.6 \pm 5	•28.1 \pm 3	•43.9 \pm 15	•12.1 \pm 5	•12.2 \pm 2	•17.9 \pm 4	•15.4 \pm 3	•7.5 \pm 4	16.9 \pm 6
LWEA (TCYB'18)	•0.4 \pm 0	•53.3 \pm 3	•15.9 \pm 3	•28.1 \pm 1	•63.3 \pm 3	•12.1 \pm 5	•13.7 \pm 3	•21.0 \pm 4	•14.7 \pm 1	•7.9 \pm 4	23.0 \pm 3
NWCA (arXiv'24)	•0.4 \pm 0	•52.5 \pm 3	•16.0 \pm 3	•28.4 \pm 1	•63.7 \pm 3	•12.5 \pm 4	•13.6 \pm 3	•21.7 \pm 1	•14.8 \pm 1	•9.7 \pm 4	23.3 \pm 2
ECCMS (TNNLS'24)	•0.4 \pm 0	•50.7 \pm 19	•18.4 \pm 5	•28.2 \pm 0	•64.7 \pm 3	•12.3 \pm 5	•12.9 \pm 3	<u>22.8</u> \pm 4	•15.5 \pm 2	•9.1 \pm 4	23.5 \pm 5
MKKM (arXiv'18)	•8.1 \pm 12	•40.8 \pm 20	•12.8 \pm 3	•20.6 \pm 6	•55.4 \pm 9	•12.0 \pm 5	•19.7 \pm 4	•14.3 \pm 4	•12.0 \pm 7	•9.1 \pm 6	20.5 \pm 8
SMKKM (TPAMI'23)	•8.7 \pm 4	•38.5 \pm 11	•19.3 \pm 4	•27.0 \pm 2	•59.4 \pm 9	•10.5 \pm 5	<u>20.0</u> \pm 2	•18.2 \pm 3	•15.5 \pm 2	•10.5 \pm 4	22.8 \pm 5
SEC (TKDE'17)	•9.2 \pm 12	•24.9 \pm 18	•17.3 \pm 4	•21.9 \pm 5	•36.0 \pm 17	•12.8 \pm 4	•15.5 \pm 3	•13.6 \pm 7	•9.9 \pm 6	•7.1 \pm 4	16.8 \pm 9
Proposed ($\alpha = 0.1$)	<u>25.0</u> \pm 12	<u>58.3</u> \pm 1	<u>20.0</u> \pm 4	<u>29.4</u> \pm 2	<u>67.5</u> \pm 3	<u>14.4</u> \pm 4	18.8 \pm 2	19.6 \pm 6	15.0 \pm 4	<u>12.4</u> \pm 4	<u>28.0</u> \pm 4
Proposed	25.0 \pm 12	59.0 \pm 1	21.1 \pm 3	29.4 \pm 2	67.5 \pm 3	15.0 \pm 4	22.9 \pm 2	27.4 \pm 2	15.8 \pm 3	12.4 \pm 4	29.6 \pm 4

Table B5. Performance (%) evaluation of different datasets based on the ARI metric. We have highlighted the values of the best-performing method in **bold**, and the second-best method is marked with an underline. • indicates whether proposed method is statistically superior to the compared methods according to the pairwise t-test at 0.05 significance level.

Method	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	Average
CEAM (TKDE'24)	•6.6 \pm 12	•42.8 \pm 31	•12.9 \pm 4	•20.4 \pm 1	•59.0 \pm 13	•2.7 \pm 5	•2.5 \pm 1	•10.8 \pm 4	•12.8 \pm 5	•10.1 \pm 7	18.1 \pm 8
CES ² L (AIJ'19)	•2.4 \pm 4	•3.0 \pm 10	•14.0 \pm 3	•20.3 \pm 2	•33.3 \pm 19	•18.3 \pm 6	•0.2 \pm 2	•6.8 \pm 7	•15.4 \pm 4	•9.6 \pm 9	12.3 \pm 7
CES ² Q (AIJ'19)	•1.7 \pm 3	•3.5 \pm 7	•12.4 \pm 3	•20.0 \pm 2	•31.2 \pm 17	•18.5 \pm 6	•0.3 \pm 2	•9.0 \pm 4	•15.2 \pm 5	•6.7 \pm 5	11.8 \pm 6
LWEA (TCYB'18)	•-0.5 \pm 0	•62.9 \pm 4	•13.1 \pm 3	•21.2 \pm 1	•57.5 \pm 5	•18.5 \pm 6	•0.0 \pm 2	•10.0 \pm 4	•13.5 \pm 3	•8.8 \pm 6	20.5 \pm 4
NWCA (arXiv'24)	•-0.5 \pm 0	•62.3 \pm 4	•12.9 \pm 2	21.6 \pm 1	•56.3 \pm 6	•19.8 \pm 5	•-0.1 \pm 2	•10.4 \pm 1	•13.3 \pm 3	•11.7 \pm 6	20.8 \pm 3
ECCMS (TNNLS'24)	•-0.5 \pm 0	•56.1 \pm 24	•13.5 \pm 3	•21.3 \pm 1	•60.8 \pm 7	•19.0 \pm 6	•-0.3 \pm 1	<u>12.2</u> \pm 4	•14.0 \pm 3	•10.5 \pm 6	20.7 \pm 6
MKKM (arXiv'18)	•8.8 \pm 14	•47.1 \pm 25	•9.5 \pm 2	•14.2 \pm 5	•53.8 \pm 10	•13.6 \pm 12	•2.1 \pm 2	•7.2 \pm 3	•10.9 \pm 6	•10.1 \pm 7	17.7 \pm 8
SMKKM (TPAMI'23)	•8.8 \pm 5	•41.9 \pm 10	•14.6 \pm 3	•17.0 \pm 3	•55.5 \pm 11	•13.2 \pm 9	<u>3.5</u> \pm 1	•7.2 \pm 4	<u>15.7</u> \pm 2	•12.2 \pm 5	19.0 \pm 5
SEC (TKDE'17)	•8.9 \pm 15	•23.8 \pm 25	•12.8 \pm 4	•13.5 \pm 5	•26.9 \pm 19	•13.5 \pm 12	•1.1 \pm 2	•5.6 \pm 7	•7.2 \pm 6	•5.2 \pm 5	11.9 \pm 9
Fix $\alpha = 0.1$	<u>30.8</u> \pm 15	<u>69.2</u> \pm 1	<u>15.8</u> \pm 4	<u>22.1</u> \pm 2	<u>67.5</u> \pm 5	<u>20.6</u> \pm 5	2.6 \pm 1	12.0 \pm 5	14.8 \pm 5	<u>14.5</u> \pm 6	<u>27.0</u> \pm 4
Proposed	30.8 \pm 15	69.5 \pm 2	16.7 \pm 3	22.1 \pm 2	67.5 \pm 5	21.5 \pm 5	4.1 \pm 1	18.4 \pm 2	16.0 \pm 3	14.5 \pm 6	28.1 \pm 3

Table B6. Performance (%) evaluation of different datasets based on the F-score metric. We have highlighted the values of the best-performing method in **bold**, and the second-best method is marked with an underline. • indicates whether proposed method is statistically superior to the compared methods according to the pairwise t-test at 0.05 significance level.

Method	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	Average
CEAM (TKDE'24)	•60.5 \pm 9	•79.4 \pm 15	•46.4 \pm 3	•42.7 \pm 1	•83.1 \pm 7	•66.0 \pm 2	•22.8 \pm 2	•51.4 \pm 3	•49.9 \pm 3	•61.8 \pm 6	56.4 \pm 5
CEs ² L (AIJ'19)	•58.0 \pm 4	•59.5 \pm 7	•46.3 \pm 2	•42.0 \pm 2	•65.0 \pm 12	•72.2 \pm 3	•19.3 \pm 2	•49.1 \pm 5	51.7 \pm 3	•62.7 \pm 6	52.6 \pm 4
CEs ² Q (AIJ'19)	•57.4 \pm 3	•60.3 \pm 5	•44.7 \pm 3	•41.9 \pm 2	•62.9 \pm 12	•72.4 \pm 3	•19.2 \pm 1	•50.5 \pm 5	•51.6 \pm 3	•60.3 \pm 4	52.1 \pm 4
LWEA (TCYB'18)	•55.5 \pm 0	•89.6 \pm 1	•46.0 \pm 3	43.2 \pm 1	•81.7 \pm 4	•72.4 \pm 3	•18.6 \pm 2	•49.5 \pm 1	•51.3 \pm 2	•61.2 \pm 4	56.9 \pm 2
NWCA (arXiv'24)	•55.5 \pm 0	•89.4 \pm 1	•45.9 \pm 2	•43.6 \pm 1	•80.7 \pm 5	73.2 \pm 2	•18.8 \pm 2	•49.2 \pm 1	•51.2 \pm 2	•63.5 \pm 4	57.1 \pm 2
ECCMS (TNNLS'24)	•55.5 \pm 0	•85.6 \pm 12	•46.1 \pm 3	•43.3 \pm 1	•84.0 \pm 3	72.6 \pm 3	•18.5 \pm 2	•51.0 \pm 3	•51.6 \pm 3	•62.5 \pm 4	57.1 \pm 3
MKKM (arXiv'18)	•62.1 \pm 10	•82.6 \pm 11	•42.9 \pm 3	•37.4 \pm 5	•79.8 \pm 7	•70.8 \pm 5	• <u>25.2</u> \pm 3	•50.2 \pm 2	•49.7 \pm 6	•62.5 \pm 6	56.3 \pm 6
SMKKM (TPAMI'23)	•62.9 \pm 4	•73.7 \pm 7	•47.7 \pm 3	•39.8 \pm 2	•80.6 \pm 8	•69.9 \pm 4	•23.4 \pm 3	•53.2 \pm 1	<u>52.2</u> \pm 1	•63.3 \pm 4	56.7 \pm 4
SEC (TKDE'17)	•62.2 \pm 10	•71.9 \pm 12	•46.0 \pm 3	•37.2 \pm 4	•59.9 \pm 13	•71.0 \pm 4	•20.5 \pm 2	•48.2 \pm 5	•45.7 \pm 5	•58.8 \pm 5	52.1 \pm 6
Fix $\alpha = 0.1$	<u>76.5</u> \pm 9	<u>91.6</u> \pm 0	<u>48.9</u> \pm 3	<u>43.7</u> \pm 1	<u>87.6</u> \pm 2	<u>73.3</u> \pm 3	21.4 \pm 2	<u>55.4</u> \pm 5	51.5 \pm 4	<u>65.1</u> \pm 5	<u>61.5</u> \pm 3
Proposed	76.5 \pm 9	91.7 \pm 1	49.8 \pm 2	43.7 \pm 1	87.6 \pm 2	73.8 \pm 2	27.3 \pm 3	63.3 \pm 1	52.3 \pm 2	65.1 \pm 5	63.1 \pm 3