



## RANKING RESULTS

This appendix exposes the ranking results of each algorithm, for each database containing clustering between 1 and 9 and for noise levels being equal to 0.5, 0.15 and 0.30.

### **A.1 Results of the ranking for Backward Ranking**

This section displays the ranking results for Backward algorithm, for each database containing clustering between 1 and 9 and for noise levels being equal to 0.5, 0.15 and 0.30.

APPENDIX A. RANKING RESULTS

Table A.1: Representation of the feature ranking for Backward Elimination, for a 0.05 noise leEel and for the databases related to 1 to 9 clusters. ID represents an empirical number giEen to each feature and characterizing its position in the feature set.

<b>Backward Ranking 0.05 noise</b>										
<b>FEATURES</b>		<b>NUMBER OF CLUSTERS</b>								
<b>ID</b>	<b>NAME</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>
1	$AMP^+$	11	7	7	15	1	15	11	15	1
2	$A^+$	15	13	13	9	15	1	15	10	15
3	$D_{pic}$	5	1	1	1	2	3	5	6	2
4	$AVG$	8	4	8	5	7	6	10	7	9
5	$E$	1	10	11	6	10	11	1	1	11
6	$\dot{AMP}^+$	9	3	5	2	4	4	6	5	3
7	$\dot{A}^+$	14	15	15	13	14	14	14	14	14
8	$\dot{D}_{pic}$	2	11	10	11	11	9	8	11	10
9	$\dot{AVG}$	12	12	12	12	12	12	12	12	12
10	$\dot{E}$	4	5	4	7	5	2	2	2	5
11	$\ddot{AMP}^+$	13	2	14	4	6	13	13	13	13
12	$\ddot{A}^+$	3	14	2	14	13	7	4	3	6
13	$\ddot{D}_{pic}$	10	9	9	3	3	5	3	4	8
14	$\ddot{AVG}$	7	6	6	8	9	10	7	9	4
15	$\ddot{E}$	6	8	3	10	8	8	9	8	7

Table A.2: Representation of the feature ranking for Backward Elimination, for a 0.15 noise leEel and for the databases related to 1 to 9 clusters. ID represents an empirical number giEen to each feature and characterizing its position in the feature set.

<b>Backward Ranking 0.15 noise</b>										
<b>FEATURES</b>		<b>NUMBER OF CLUSTERS</b>								
<b>ID</b>	<b>NAME</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>
1	$AMP^+$	11	12	11	6	1	1	15	15	3
2	$A^+$	13	15	14	15	15	15	9	3	15
3	$D_{pic}$	10	1	15	11	9	11	11	6	10
4	$AVG$	8	6	8	3	6	3	10	11	7
5	$E$	6	10	3	1	7	9	1	9	12
6	$\dot{A}MP^+$	12	13	4	8	4	7	13	1	1
7	$\dot{A}^+$	15	2	12	14	14	14	5	14	14
8	$\dot{D}_{pic}$	5	3	6	5	8	8	4	4	8
9	$\dot{A}VG$	9	11	1	12	12	12	12	12	11
10	$\dot{E}$	3	7	5	4	5	4	3	7	6
11	$\ddot{A}MP^+$	14	14	2	13	3	2	8	5	4
12	$\ddot{A}^+$	1	5	13	2	13	13	14	13	13
13	$\ddot{D}_{pic}$	4	9	9	7	11	6	7	8	5
14	$\ddot{A}VG$	7	8	10	10	2	5	2	2	2
15	$\ddot{E}$	2	4	7	9	10	10	6	10	9

Table A.3: Representation of the feature ranking for Backward Elimination, for a 0.30 noise leEel and for the databases related to 1 to 9 clusters. ID represents an empirical number giEen to each feature and characterizing its position in the feature set.

<b>Backward Ranking 0.3 noise</b>										
<b>FEATURES</b>		<b>NUMBER OF CLUSTERS</b>								
<b>ID</b>	<b>NAME</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>
1	$AMP^+$	14	15	12	3	8	15	10	4	15
2	$A^+$	11	12	13	15	15	11	15	15	1
3	$D_{pic}$	10	11	14	10	13	9	6	1	7
4	$AVG$	1	6	1	7	10	12	11	14	13
5	$E$	5	9	11	9	11	1	1	11	14
6	$\dot{A}MP^+$	13	8	5	1	3	4	9	6	3
7	$\dot{A}^+$	9	1	9	13	12	13	7	12	12
8	$\dot{D}_{pic}$	15	3	6	4	6	6	3	3	8
9	$\dot{A}VG$	3	2	2	11	5	2	14	10	2
10	$\dot{E}$	6	4	4	5	4	5	4	7	6
11	$\ddot{A}MP^+$	4	5	10	6	9	10	12	2	5
12	$\ddot{A}^+$	12	13	3	12	1	3	2	13	10
13	$\ddot{D}_{pic}$	8	14	15	14	14	14	13	8	11
14	$\ddot{A}VG$	7	10	8	2	2	8	5	5	4
15	$\ddot{E}$	2	7	7	8	7	7	8	9	9

## **A.2 Results of the ranking for Feature Dispersion Ranking**

This section displays the ranking results for Feature Dispersion algorithm, for each database containing clustering between 1 and 9 and for noise levels being equal to 0.5, 0.15 and 0.30.

Table A.4: Representation of the feature ranking for Feature Dispersion, for a 0.05 noise level and for the databases related to 1 to 9 clusters. ID represents an empirical number given to each feature and characterizing its position in the feature set.

Feature Dispersion Ranking 0.05 noise										
FEATURES		NUMBER OF CLUSTERS								
ID	NAME	1	2	3	4	5	6	7	8	9
1	$AMP^+$	9	9	11	8	7	9	11	14	10
2	$A^+$	10	10	12	9	8	10	12	15	11
3	$D_{pic}$	3	3	3	3	3	3	3	3	3
4	$AVG$	5	5	5	6	6	4	6	5	6
5	$E$	7	13	4	7	12	11	10	7	9
6	$\dot{AMP}^+$	12	14	8	12	13	12	14	11	14
7	$\dot{A}^+$	13	15	9	13	14	13	15	12	15
8	$\dot{D}_{pic}$	1	1	1	1	1	1	1	1	1
9	$\dot{AVG}$	4	6	6	4	4	5	5	4	4
10	$\dot{E}$	11	8	10	15	15	15	9	13	12
11	$\ddot{AMP}^+$	14	11	13	10	11	7	7	8	7
12	$\ddot{A}^+$	15	12	14	11	10	8	8	9	8
13	$\ddot{D}_{pic}$	2	2	2	2	2	2	2	2	2
14	$\ddot{AVG}$	6	7	7	5	5	6	4	6	5
15	$\ddot{E}$	8	4	15	14	9	14	13	10	13

Table A.5: Representation of the feature ranking for Feature Dispersion, for a 0.15 noise level and for the databases related to 1 to 9 clusters. ID represents an empirical number given to each feature and characterizing its position in the feature set.

Feature Dispersion Ranking 0.15 noise										
FEATURES		NUMBER OF CLUSTERS								
ID	NAME	1	2	3	4	5	6	7	8	9
1	$AMP^+$	7	13	7	8	8	10	8	9	12
2	$A^+$	8	14	8	9	9	11	9	10	13
3	$D_{pic}$	2	3	9	3	3	3	3	3	3
4	$AVG$	1	6	5	6	6	6	6	6	6
5	$E$	5	12	6	7	7	13	7	13	9
6	$\dot{AMP}^+$	9	10	10	12	13	14	12	12	14
7	$\dot{A}^+$	10	11	11	13	14	15	13	7	15
8	$\dot{D}_{pic}$	3	1	1	1	1	1	1	1	1
9	$\dot{AVG}$	4	5	4	4	4	4	4	4	4
10	$\dot{E}$	12	15	12	14	10	12	14	11	10
11	$\ddot{AMP}^+$	15	7	14	10	11	8	10	14	7
12	$\ddot{A}^+$	14	8	15	11	12	9	11	15	8
13	$\ddot{D}_{pic}$	13	2	2	2	2	2	2	2	2
14	$\ddot{AVG}$	6	4	3	5	5	5	5	5	5
15	$\ddot{E}$	11	9	13	15	15	7	15	8	11

Table A.6: Representation of the feature ranking for Feature Dispersion, for a 0.30 noise level and for the databases related to 1 to 9 clusters. ID represents an empirical number given to each feature and characterizing its position in the feature set.

Feature Dispersion Ranking 0.3 noise										
FEATURES		NUMBER OF CLUSTERS								
ID	NAME	1	2	3	4	5	6	7	8	9
1	$AMP^+$	5	7	8	14	12	9	14	10	8
2	$A^+$	6	8	9	15	13	10	13	11	9
3	$D_{pic}$	1	3	15	3	3	2	3	3	3
4	$AVG$	2	12	2	5	6	6	4	6	5
5	$E$	7	13	10	7	10	15	11	7	7
6	$\dot{A}MP^+$	8	5	5	9	14	13	9	14	12
7	$\dot{A}^+$	9	4	4	10	15	14	10	15	13
8	$\dot{D}_{pic}$	11	2	1	1	1	1	1	1	1
9	$\dot{A}VG$	15	1	3	4	2	3	6	5	4
10	$\dot{E}$	4	6	7	13	9	11	12	12	15
11	$\ddot{A}MP^+$	13	15	12	11	7	7	7	8	10
12	$\ddot{A}^+$	14	14	13	12	8	8	8	9	11
13	$\ddot{D}_{pic}$	3	10	14	2	5	4	2	2	2
14	$\ddot{A}VG$	12	11	6	6	4	5	5	4	6
15	$\ddot{E}$	10	9	11	8	11	12	15	13	14

### **A.3 Results of the ranking for Forward1 Ranking**

This section displays the ranking results for Forward Selection 1 algorithm, for each database containing clustering between 1 and 9 and for noise levels being equal to 0.5, 0.15 and 0.30.

APPENDIX A. RANKING RESULTS

Table A.7: Representation of the feature ranking for Forward Selection 1, for a 0.05 noise leEel and for the databases related to 1 to 9 clusters. ID represents an empirical number giEen to each feature and characterizing its position in the feature set.

Forward1 Ranking 0.05 noise										
FEATURES		NUMBER OF CLUSTERS								
ID	NAME	1	2	3	4	5	6	7	8	9
1	$AMP^+$	11	2	2	4	2	7	2	15	15
2	$A^+$	15	13	13	15	15	15	15	2	2
3	$D_{pic}$	1	1	4	1	1	5	1	1	1
4	$AVG$	8	4	9	6	9	2	11	8	9
5	$E$	3	10	11	7	10	11	8	9	11
6	$\dot{A}MP^+$	14	3	6	5	4	3	3	3	3
7	$\dot{A}^+$	9	15	15	13	14	14	14	14	14
8	$\dot{D}_{pic}$	7	11	5	11	11	9	9	11	10
9	$\dot{A}VG$	12	12	12	12	12	12	12	12	12
10	$\dot{E}$	6	5	1	8	5	1	5	5	6
11	$\ddot{A}MP^+$	4	14	14	2	6	4	13	6	7
12	$\ddot{A}^+$	13	6	7	14	13	13	6	13	13
13	$\ddot{D}_{pic}$	10	9	10	3	3	6	4	4	4
14	$\ddot{A}VG$	2	7	8	9	7	10	7	10	5
15	$\ddot{E}$	5	8	3	10	8	8	10	7	8

Table A.8: Representation of the feature ranking for Forward Selection 1, for a 0.15 noise leEel and for the databases related to 1 to 9 clusters. ID represents an empirical number giEen to each feature and characterizing its position in the feature set.

<b>Forward1 Ranking 0.15 noise</b>										
<b>FEATURES</b>		<b>NUMBER OF CLUSTERS</b>								
<b>ID</b>	<b>NAME</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>
1	$AMP^+$	13	2	11	2	2	15	2	9	4
2	$A^+$	12	15	14	15	15	5	15	15	15
3	$D_{pic}$	1	9	15	1	9	11	1	4	10
4	$AVG$	10	6	8	9	7	7	11	11	7
5	$E$	3	12	2	6	8	9	6	3	12
6	$\dot{A}MP^+$	15	4	3	3	3	14	13	14	2
7	$\dot{A}^+$	2	13	12	14	14	2	8	2	14
8	$\dot{D}_{pic}$	8	8	1	8	10	8	7	1	8
9	$\dot{A}VG$	7	11	4	12	12	12	12	12	11
10	$\dot{E}$	4	5	6	5	6	6	4	6	6
11	$\ddot{A}MP^+$	11	10	13	7	13	4	3	5	5
12	$\ddot{A}^+$	14	14	5	13	5	13	14	13	13
13	$\ddot{D}_{pic}$	5	1	9	10	1	1	10	8	1
14	$\ddot{A}VG$	9	3	10	4	4	3	5	14	3
15	$\ddot{E}$	6	7	7	11	11	10	9	10	9

Table A.9: Representation of the feature ranking for Forward Selection 1, for a 0.30 noise leEel and for the databases related to 1 to 9 clusters. ID represents an empirical number giEen to each feature and characterizing its position in the feature set.

Forward1 Ranking 0.3 noise										
FEATURES		NUMBER OF CLUSTERS								
ID	NAME	1	2	3	4	5	6	7	8	9
1	$AMP^+$	2	12	13	15	15	11	11	15	8
2	$A^+$	14	15	12	9	11	15	15	9	15
3	$D_{pic}$	11	11	14	10	13	9	9	8	9
4	$AVG$	9	7	2	5	10	12	13	14	13
5	$E$	6	9	11	4	8	6	7	11	14
6	$\dot{A}MP^+$	10	3	5	13	3	13	14	2	2
7	$\dot{A}^+$	13	8	9	7	12	3	8	12	12
8	$\dot{D}_{pic}$	15	4	6	1	6	5	5	1	5
9	$\dot{A}VG$	7	2	1	11	5	8	2	10	7
10	$\dot{E}$	5	1	4	2	4	4	6	6	4
11	$\ddot{A}MP^+$	12	6	10	12	2	2	12	5	3
12	$\ddot{A}^+$	4	13	3	3	9	10	3	13	10
13	$\ddot{D}_{pic}$	1	14	15	14	14	14	1	4	11
14	$\ddot{A}VG$	8	10	8	8	1	1	4	3	1
15	$\ddot{E}$	3	5	7	6	7	7	10	7	6

## **A.4 Results of the ranking for Forward2 Ranking**

This section displays the ranking results for Forward Selection 2 algorithm, for each database containing clustering between 1 and 9 and for noise levels being equal to 0.5, 0.15 and 0.30.

APPENDIX A. RANKING RESULTS

Table A.10: Representation of the feature ranking for Forward Selection 2, for a 0.05 noise leEel and for the databases related to 1 to 9 clusters. ID represents an empirical number giEen to each feature and characterizing its position in the feature set.

Forward2 Ranking 0.05 noise										
FEATURES		NUMBER OF CLUSTERS								
ID	NAME	1	2	3	4	5	6	7	8	9
1	$AMP^+$	14	10	10	14	12	14	12	14	13
2	$A^+$	15	11	11	15	13	15	13	15	12
3	$D_{pic}$	1	1	2	1	1	2	1	1	1
4	$AVG$	3	6	6	3	5	4	4	6	4
5	$E$	7	8	9	5	8	9	8	8	9
6	$\dot{A}MP^+$	13	14	14	10	14	12	14	13	14
7	$\dot{A}^+$	12	15	15	11	15	13	15	12	15
8	$\dot{D}_{pic}$	5	5	8	7	6	7	5	9	6
9	$\dot{A}VG$	8	7	7	6	7	6	7	7	7
10	$\dot{E}$	9	4	1	8	2	1	2	2	3
11	$\ddot{A}MP^+$	10	12	12	13	10	10	11	10	11
12	$\ddot{A}^+$	11	13	13	12	11	11	10	11	10
13	$\ddot{D}_{pic}$	2	2	3	2	3	3	3	3	2
14	$\ddot{A}VG$	6	3	5	4	4	5	6	4	5
15	$\ddot{E}$	4	9	4	9	9	8	9	5	8

Table A.11: Representation of the feature ranking for Forward Selection 2, for a 0.15 noise leEel and for the databases related to 1 to 9 clusters. ID represents an empirical number giEen to each feature and characterizing its position in the feature set.

<b>Forward2 Ranking 0.15 noise</b>										
<b>FEATURES</b>		<b>NUMBER OF CLUSTERS</b>								
<b>ID</b>	<b>NAME</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>
1	<i>AMP</i> <sup>+</sup>	10	14	10	14	14	14	14	14	14
2	<i>A</i> <sup>+</sup>	11	15	11	15	15	15	15	15	15
3	<i>D<sub>pic</sub></i>	1	2	7	1	3	2	1	2	3
4	<i>AVG</i>	8	7	3	6	6	5	3	7	5
5	<i>E</i>	3	9	9	3	5	7	4	4	7
6	<i>AMP</i> <sup>+</sup>	14	11	14	13	12	13	10	12	12
7	<i>A</i> <sup>+</sup>	15	10	15	12	13	12	11	13	13
8	<i>D<sub>pic</sub></i>	7	4	1	4	8	8	6	1	8
9	<i>AVG</i>	9	6	4	7	7	6	7	8	6
10	<i>E</i>	4	3	2	2	2	3	2	3	4
11	<i>AMP</i> <sup>+</sup>	12	12	12	10	10	10	12	11	10
12	<i>A</i> <sup>+</sup>	13	13	13	11	11	11	13	10	11
13	<i>D<sub>pic</sub></i>	5	1	5	8	1	1	8	5	1
14	<i>AVG</i>	6	5	6	5	4	4	5	6	2
15	<i>E</i>	2	8	8	9	9	9	9	9	9

Table A.12: Representation of the feature ranking for Forward Selection 2, for a 0.30 noise leEel and for the databases related to 1 to 9 clusters. ID represents an empirical number giEen to each feature and characterizing its position in the feature set.

Forward2 Ranking 0.3 noise										
FEATURES		NUMBER OF CLUSTERS								
ID	NAME	1	2	3	4	5	6	7	8	9
1	$AMP^+$	9	14	14	14	14	14	14	14	14
2	$A^+$	8	15	15	15	15	15	15	15	15
3	$D_{pic}$	14	7	6	4	7	5	2	2	2
4	$AVG$	6	5	5	7	3	3	5	6	5
5	$E$	5	9	10	6	9	7	8	7	7
6	$\dot{A}MP^+$	10	10	7	11	10	12	11	10	12
7	$\dot{A}^+$	11	11	8	10	11	11	10	11	11
8	$\dot{D}_{pic}$	15	3	12	1	12	8	4	1	8
9	$\dot{A}VG$	4	2	1	8	2	2	6	5	4
10	$\dot{E}$	7	1	9	3	4	4	7	8	6
11	$\ddot{A}MP^+$	13	13	3	12	6	10	12	13	9
12	$\ddot{A}^+$	12	12	4	13	5	9	13	12	10
13	$\ddot{D}_{pic}$	1	6	11	5	8	6	1	3	3
14	$\ddot{A}VG$	2	4	2	2	1	1	3	4	1
15	$\ddot{E}$	3	8	13	9	13	13	9	9	15

## **A.5 Results of the ranking for KDE Ranking**

This section displays the ranking results for Kernel Density Estimation 1 algorithm, for each database containing clustering between 1 and 9 and for noise levels being equal to 0.5, 0.15 and 0.30.

Table A.13: Representation of the feature ranking for Kernel Density Estimation, for a 0.05 noise level and for the databases related to 1 to 9 clusters. ID represents an empirical number given to each feature and characterizing its position in the feature set.

<b>KDE Ranking 0.05 noise</b>										
<b>FEATURES</b>		<b>NUMBER OF CLUSTERS</b>								
<b>ID</b>	<b>NAME</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>
1	$AMP^+$	14	5	4	8	10	8	11	13	8
2	$A^+$	15	4	5	9	9	9	10	12	7
3	$D_{pic}$	3	6	6	4	3	3	3	3	4
4	$AVG$	6	7	7	7	6	4	9	5	13
5	$E$	13	10	3	10	7	7	7	11	10
6	$\dot{AMP}^+$	11	8	10	11	13	12	14	8	15
7	$\dot{A}^+$	12	9	9	12	12	11	15	7	14
8	$\dot{D}_{pic}$	2	1	1	1	1	2	2	2	2
9	$\dot{AVG}$	4	11	12	2	4	6	8	4	5
10	$\dot{E}$	5	13	11	6	8	13	6	10	9
11	$\ddot{AMP}^+$	7	15	14	15	14	14	12	15	11
12	$\ddot{A}^+$	8	14	15	14	15	15	13	14	12
13	$\ddot{D}_{pic}$	1	2	2	13	2	1	1	1	1
14	$\ddot{AVG}$	9	12	13	3	11	10	5	9	6
15	$\ddot{E}$	10	3	8	5	5	5	4	6	3

Table A.14: Representation of the feature ranking for Kernel Density Estimation, for a 0.15 noise leEel and for the databases related to 1 to 9 clusters. ID represents an empirical number giEen to each feature and characterizing its position in the feature set.

<b>KDE Ranking 0.15 noise</b>										
<b>FEATURES</b>		<b>NUMBER OF CLUSTERS</b>								
<b>ID</b>	<b>NAME</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>
1	$AMP^+$	4	9	2	12	10	7	7	8	14
2	$A^+$	3	8	3	13	9	6	6	7	15
3	$D_{pic}$	2	6	11	3	2	2	4	1	4
4	$AVG$	1	11	10	5	6	8	13	6	8
5	$E$	13	13	13	7	11	10	8	9	13
6	$\dot{A}MP^+$	14	2	7	15	15	12	11	14	12
7	$\dot{A}^+$	15	3	6	14	14	13	12	15	11
8	$\dot{D}_{pic}$	7	5	9	10	7	9	9	10	1
9	$\dot{A}VG$	5	4	5	4	3	4	3	3	6
10	$\dot{E}$	6	12	12	6	4	3	5	5	5
11	$\ddot{A}MP^+$	10	14	15	8	12	14	14	13	10
12	$\ddot{A}^+$	9	15	14	9	13	15	15	12	9
13	$\ddot{D}_{pic}$	8	1	1	11	8	11	10	11	2
14	$\ddot{A}VG$	11	10	8	2	5	5	1	4	7
15	$\ddot{E}$	12	7	4	1	1	1	2	2	3

Table A.15: Representation of the feature ranking for Kernel Density Estimation, for a 0.30 noise leEel and for the databases related to 1 to 9 clusters. ID represents an empirical number giEen to each feature and characterizing its position in the feature set.

<b>KDE Ranking 0.3 noise</b>										
<b>FEATURES</b>		<b>NUMBER OF CLUSTERS</b>								
<b>ID</b>	<b>NAME</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>
1	$AMP^+$	14	15	14	14	14	14	14	11	13
2	$A^+$	15	14	15	13	15	15	15	10	14
3	$D_{pic}$	1	1	12	2	6	3	3	2	5
4	$AVG$	6	10	7	7	13	6	6	7	7
5	$E$	4	7	13	6	7	7	5	5	4
6	$\dot{A}MP^+$	8	6	4	10	11	12	7	8	8
7	$\dot{A}^+$	7	5	3	11	10	11	8	9	9
8	$\dot{D}_{pic}$	2	11	10	12	12	10	12	12	12
9	$\dot{A}\dot{V}G$	11	3	6	3	4	4	9	6	6
10	$\dot{E}$	3	8	5	4	3	2	4	4	2
11	$\ddot{A}MP^+$	9	12	8	8	8	9	10	14	10
12	$\ddot{A}^+$	10	13	9	9	9	8	11	15	11
13	$\ddot{D}_{pic}$	5	2	2	15	2	13	13	13	15
14	$\ddot{A}\ddot{V}G$	13	9	11	5	5	5	2	3	3
15	$\ddot{E}$	12	4	1	1	1	1	1	1	1

## **A.6 Results of the ranking for Simple Ranking**

This section displays the ranking results for Simple Ranking 1 algorithm, for each database containing clustering between 1 and 9 and for noise levels being equal to 0.5, 0.15 and 0.30.

Table A.16: Representation of the feature ranking for Simple Ranking, for a 0.05 noise level and for the databases related to 1 to 9 clusters. ID represents an empirical number given to each feature and characterizing its position in the feature set.

<b>Simple Ranking 0.05 noise</b>										
<b>FEATURES</b>		<b>NUMBER OF CLUSTERS</b>								
<b>ID</b>	<b>NAME</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>
1	$AMP^+$	14	11	11	15	14	15	14	15	14
2	$A^+$	15	10	10	14	15	14	15	14	15
3	$D_{pic}$	1	1	2	1	1	2	1	1	1
4	$AVG$	3	5	6	3	5	4	4	7	5
5	$E$	7	7	8	4	7	8	6	6	8
6	$\dot{A}MP^+$	12	14	14	10	13	12	13	13	12
7	$\dot{A}^+$	13	15	15	11	12	13	12	12	13
8	$\dot{D}_{pic}$	5	6	7	8	8	6	5	8	7
9	$\dot{A}VG$	9	9	9	9	9	9	9	9	9
10	$\dot{E}$	6	3	1	5	2	1	2	2	3
11	$\ddot{A}MP^+$	11	12	13	12	11	10	11	10	10
12	$\ddot{A}^+$	10	13	12	13	10	11	10	11	11
13	$\ddot{D}_{pic}$	2	2	4	2	3	3	3	3	2
14	$\ddot{A}VG$	8	4	5	6	4	5	7	5	4
15	$\ddot{E}$	4	8	3	7	6	7	8	4	6

Table A.17: Representation of the feature ranking for Simple Ranking, for a 0.15 noise level and for the databases related to 1 to 9 clusters. ID represents an empirical number given to each feature and characterizing its position in the feature set.

Simple Ranking 0.15 noise										
FEATURES		NUMBER OF CLUSTERS								
ID	NAME	1	2	3	4	5	6	7	8	9
1	$AMP^+$	11	14	13	14	14	14	15	15	14
2	$A^+$	10	15	14	15	15	15	14	14	15
3	$D_{pic}$	1	2	15	1	3	3	1	2	3
4	$AVG$	8	7	3	6	6	4	5	7	6
5	$E$	5	9	8	3	5	6	3	4	9
6	$\dot{A}MP^+$	14	11	9	12	13	13	10	13	13
7	$\dot{A}^+$	15	10	10	13	12	12	11	12	12
8	$\dot{D}_{pic}$	7	4	1	4	7	7	4	1	5
9	$\dot{A}VG$	9	8	4	9	9	8	9	9	8
10	$\dot{E}$	3	3	2	2	2	2	2	3	4
11	$\ddot{A}MP^+$	12	12	11	10	10	10	13	10	10
12	$\ddot{A}^+$	13	13	12	11	11	11	12	11	11
13	$\ddot{D}_{pic}$	4	1	5	5	1	1	7	5	1
14	$\ddot{A}VG$	6	6	6	7	4	5	6	6	2
15	$\ddot{E}$	2	5	7	8	8	9	8	8	7

Table A.18: Representation of the feature ranking for Simple Ranking, for a 0.30 noise level and for the databases related to 1 to 9 clusters. ID represents an empirical number given to each feature and characterizing its position in the feature set.

Simple Ranking 0.3 noise										
FEATURES		NUMBER OF CLUSTERS								
ID	NAME	1	2	3	4	5	6	7	8	9
1	$AMP^+$	13	15	13	14	14	15	14	14	15
2	$A^+$	14	14	14	15	15	14	15	15	14
3	$D_{pic}$	10	9	11	4	12	9	4	2	2
4	$AVG$	7	5	5	6	6	7	8	13	12
5	$E$	5	13	12	5	11	12	6	8	13
6	$\dot{A}MP^+$	8	8	7	12	8	10	11	10	10
7	$\dot{A}^+$	9	7	6	13	9	11	12	9	9
8	$\dot{D}_{pic}$	15	3	9	1	7	4	2	1	6
9	$\dot{A}VG$	4	2	1	8	2	2	13	7	3
10	$\dot{E}$	6	1	8	3	3	3	5	5	4
11	$\ddot{A}MP^+$	12	11	4	11	4	6	10	11	7
12	$\ddot{A}^+$	11	10	3	10	5	5	9	12	8
13	$\ddot{D}_{pic}$	1	12	15	9	13	13	1	3	5
14	$\ddot{A}VG$	3	4	2	2	1	1	3	4	1
15	$\ddot{E}$	2	6	10	7	10	8	7	6	11

## PERFORMANCES RESULTS

The following conventions are taken for the plots of performance values:

- For noise (symbols):
  - asterisk for noise level 0.05
  - star for noise level 0.10
  - circle for noise level 0.15
  - square for noise level 0.20
  - hexagon for noise level 0.30
  
- For algorithm (colors):
  - blue for Backward Ranking (Backward)
  - green for Feature Dispersion (FD)
  - red for Forward Selection 1 (Forward1)
  - orange for Forward Selection 2 (Forward2)
  - dark magenta for Kernel Density Evaluation (KDE)
  - pink for Principal Component Analysis (PCA)
  - brown for Simple Ranking (Simple)

## **B.1 Comparison in Performances**

This section exposes graphs comparing the performance levels between the algorithms, for database containing numbers of clusters between 2 and 9 and for a 0.15 noise level.

## B.1. COMPARISON IN PERFORMANCES

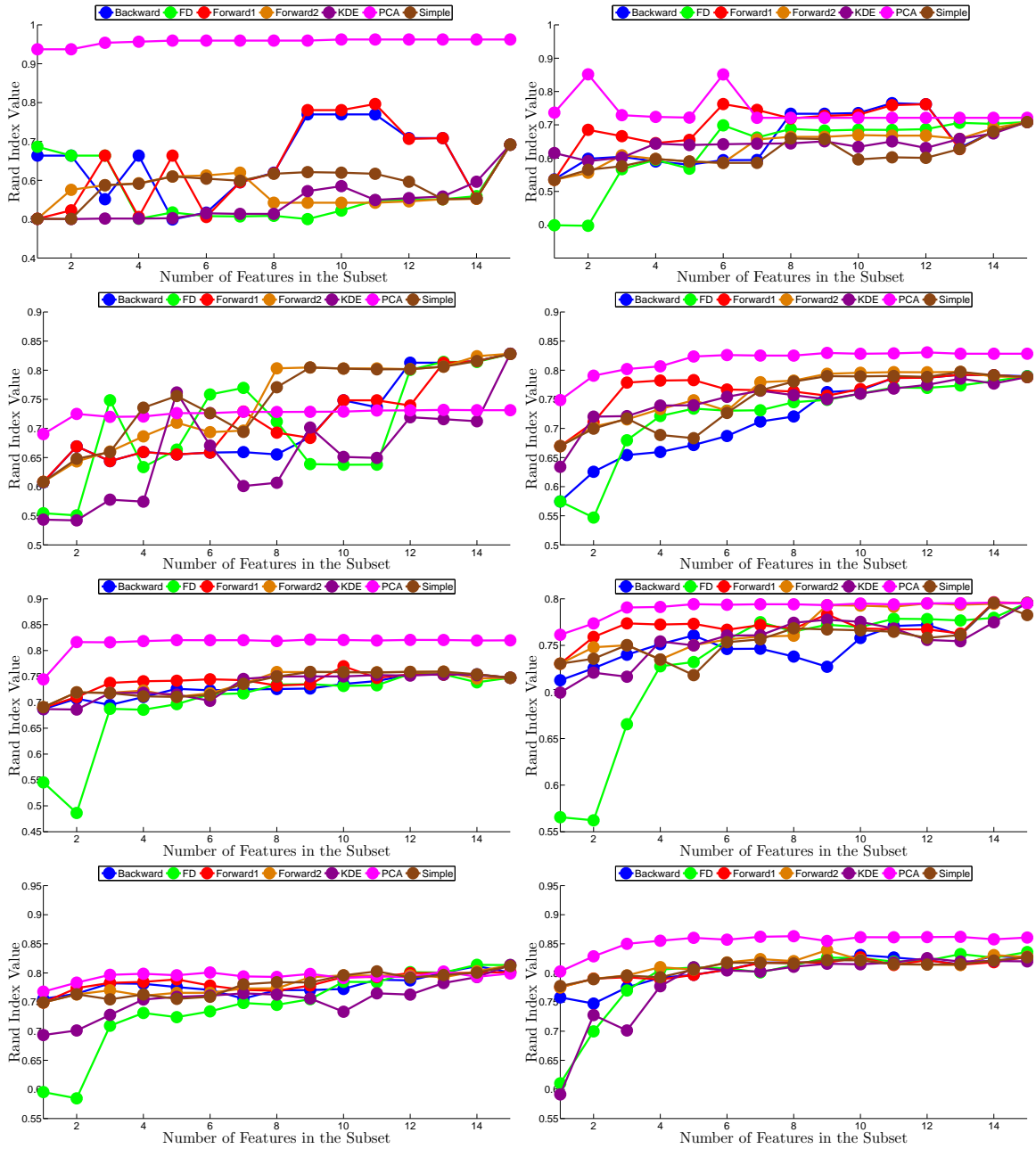


Figure B.1: Representation of the rand index for an increasing number of features, for each algorithm and for a noise level equal to 0.15. The maximum number of features is equal to 15 for each algorithm except PCA since the database does not count more than 15 features. For PCA, this stop level has been arbitrarily set to 15 too. The concordance is the following: Backward = Backward Elimination; Forward1 = Forward Selection 1; Forward2 = Forward Selection 2; Simple = Simple Ranking (for them, see subsection 4.4.1; KDE = Kernel Density Estimation of the Feature Density Similarity (see subsection 4.4.2); PCA = Principal Component Analysis. The disposition is the following: upper left = 2 clusters; upper right = 3 clusters; upper middle left = 4 clusters; upper middle right = 5 clusters; lower middle left = 6 clusters; lower middle right = 7 clusters; lower left = 8 clusters; lower right = 9 clusters.

## **B.2 Impact of noise on Algorithm Performances**

This section exposes graphs comparing the performance levels of each algorithm separately and between the algorithms, for numbers of clusters in the database equal to 3, 5 and 9 and noise levels equal to 0.05, 0.10, 0.15, 0.20 and 0.30.

## B.2. IMPACT OF NOISE ON ALGORITHM PERFORMANCES

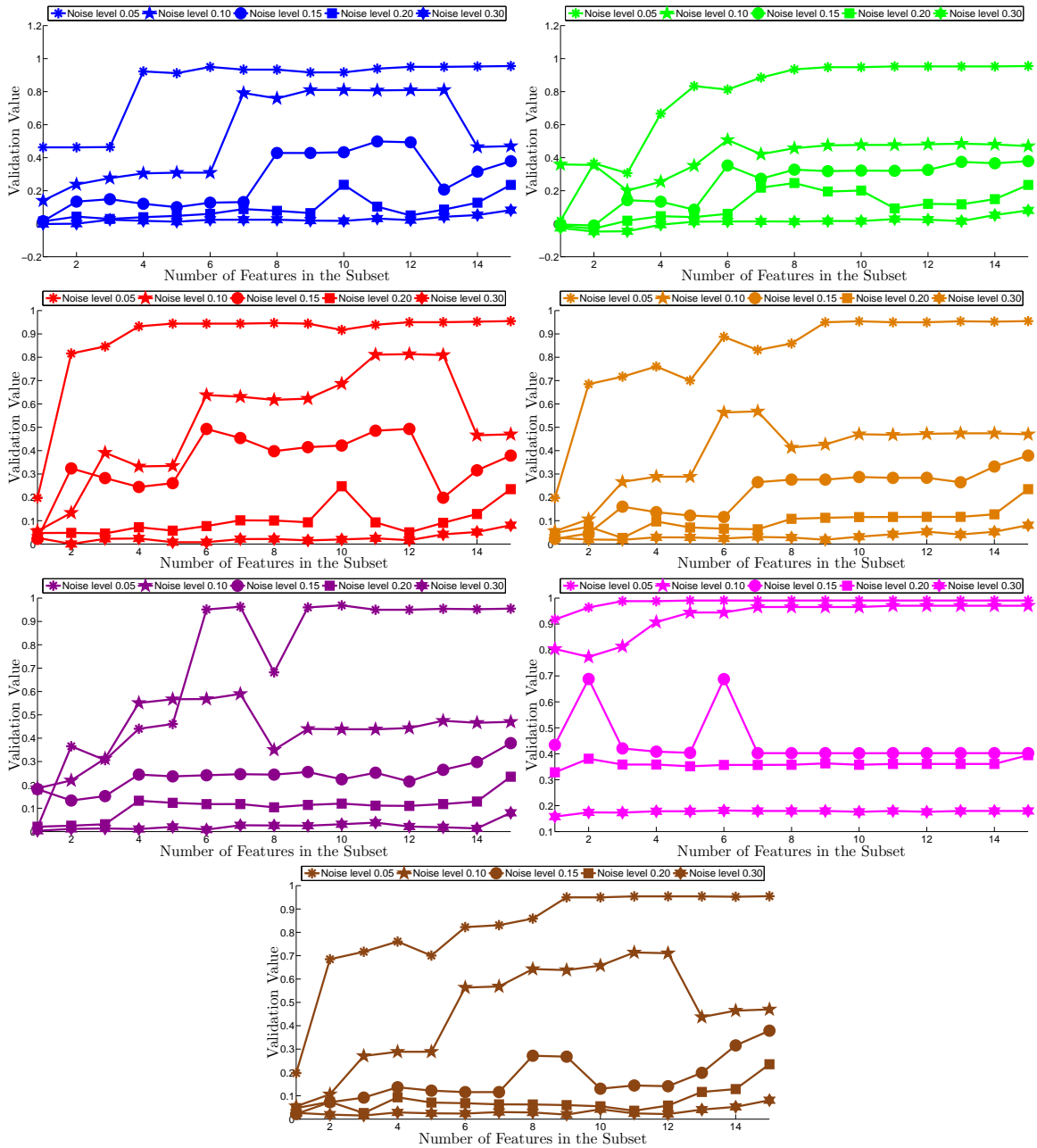


Figure B.2: Representation of the Rand Index value for the algorithms according to an increasing number of features, for 3 clusters and a noise levels being respectively equal to 0.05 (asterisk), 0.10 (star), 0.15 (circle), 0.20 (square) or 0.30 (hexagon). The concordance is the following: blue = Backward Ranking (Backward); green = Feature Dispersion (FD); red = Forward Selection 1 (Forward 1); orange = Forward Selection 2 (Forward 2); dark magenta = Kernel Density Estimation (KDE); pink = Principal Component Analysis (PCA); brown = Simple Ranking (Simple)

APPENDIX B. PERFORMANCES RESULTS

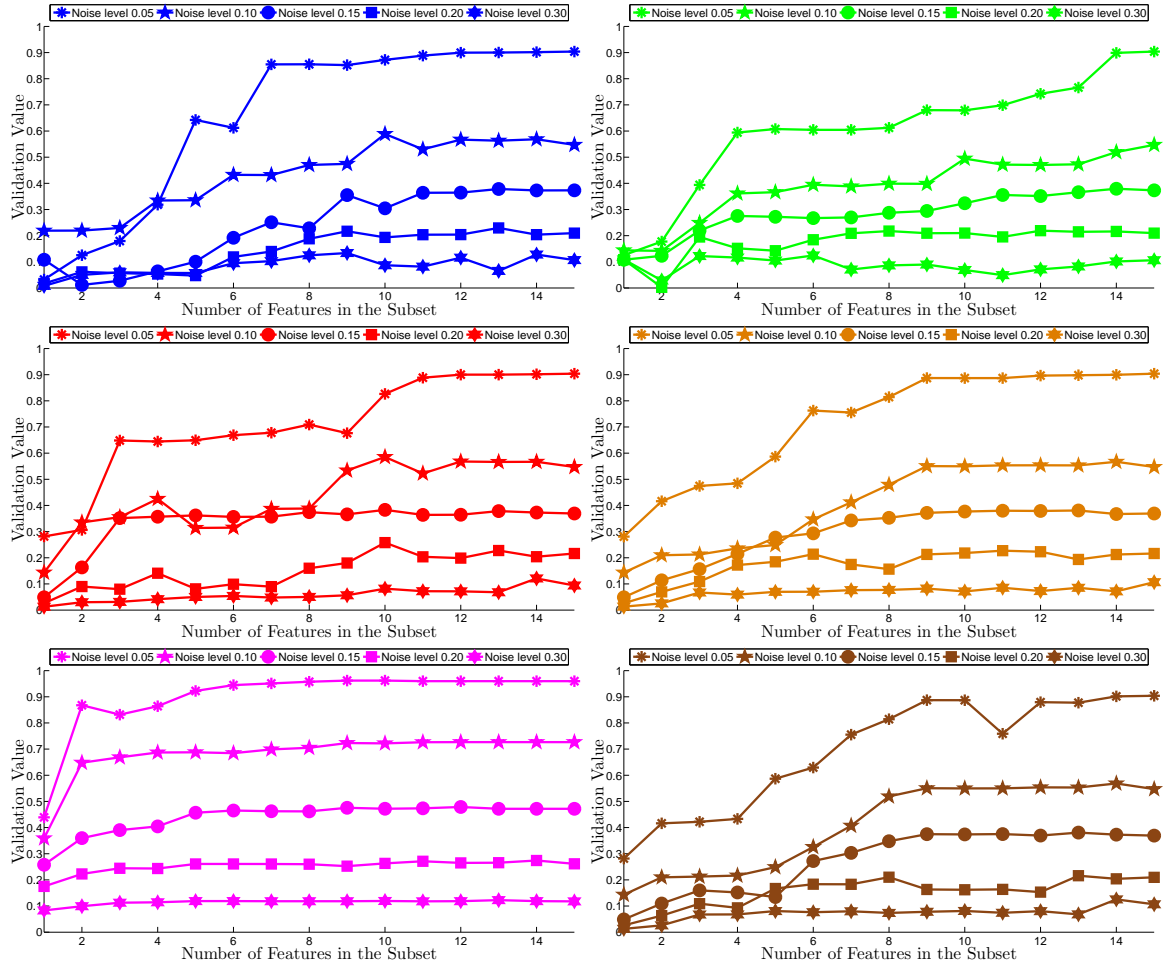


Figure B.3: Representation of the Rand Index value for the algorithms according to an increasing number of features, for 5 clusters and a noise levels being respectively equal to 0.05 (asterisk), 0.10 (star), 0.15 (circle), 0.20 (square) or 0.30 (hexagon). The concordance is the following: blue = Backward Ranking (Backward); green = Feature Dispersion (FD); red = Forward Selection 1 (Forward 1); orange = Forward Selection 2 (Forward 2); pink = Principal Component Analysis (PCA); brown = Simple Ranking (Simple)

## B.2. IMPACT OF NOISE ON ALGORITHM PERFORMANCES

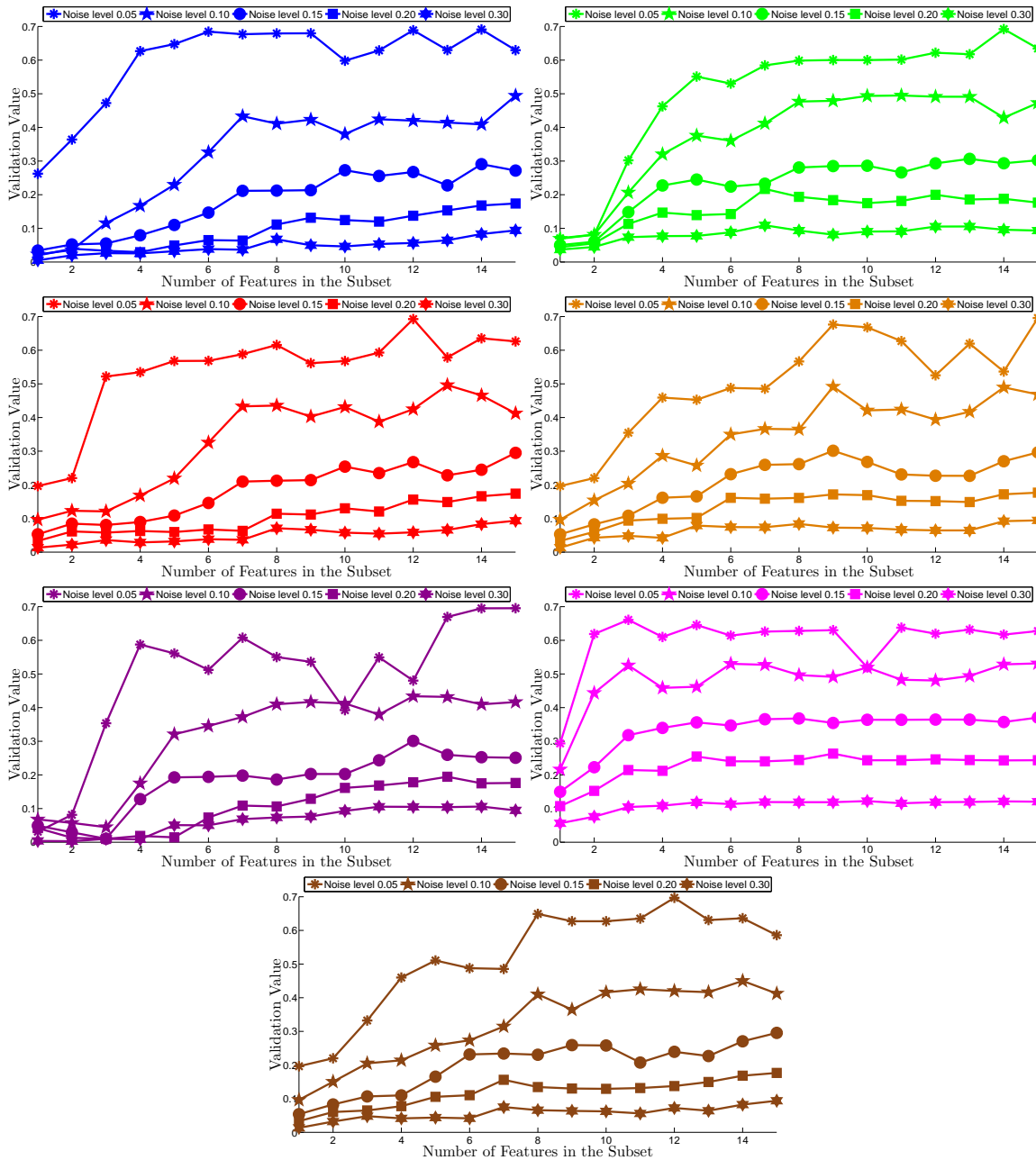


Figure B.4: Representation of the Rand Index value for the algorithms according to an increasing number of features, for 9 clusters and a noise levels being respectively equal to 0.05 (asterisk), 0.10 (star), 0.15 (circle), 0.20 (square) or 0.30 (hexagon). The concordance is the following: blue = Backward Ranking (Backward); green = Feature Dispersion (FD); red = Forward Selection 1 (Forward 1); orange = Forward Selection 2 (Forward 2); dark magenta = Kernel Density Estimation (KDE); pink = Principal Component Analysis (PCA); brown = Simple Ranking (Simple)

### **B.3 Impact of noise on Comparison of Algorithm Performances**

This section exposes graphs comparing the performance levels between the algorithms, for database containing numbers of clusters between 2 and 9 and for a 0.05 and 0.30 noise levels.

### B.3. IMPACT OF NOISE ON COMPARISON OF ALGORITHM PERFORMANCES

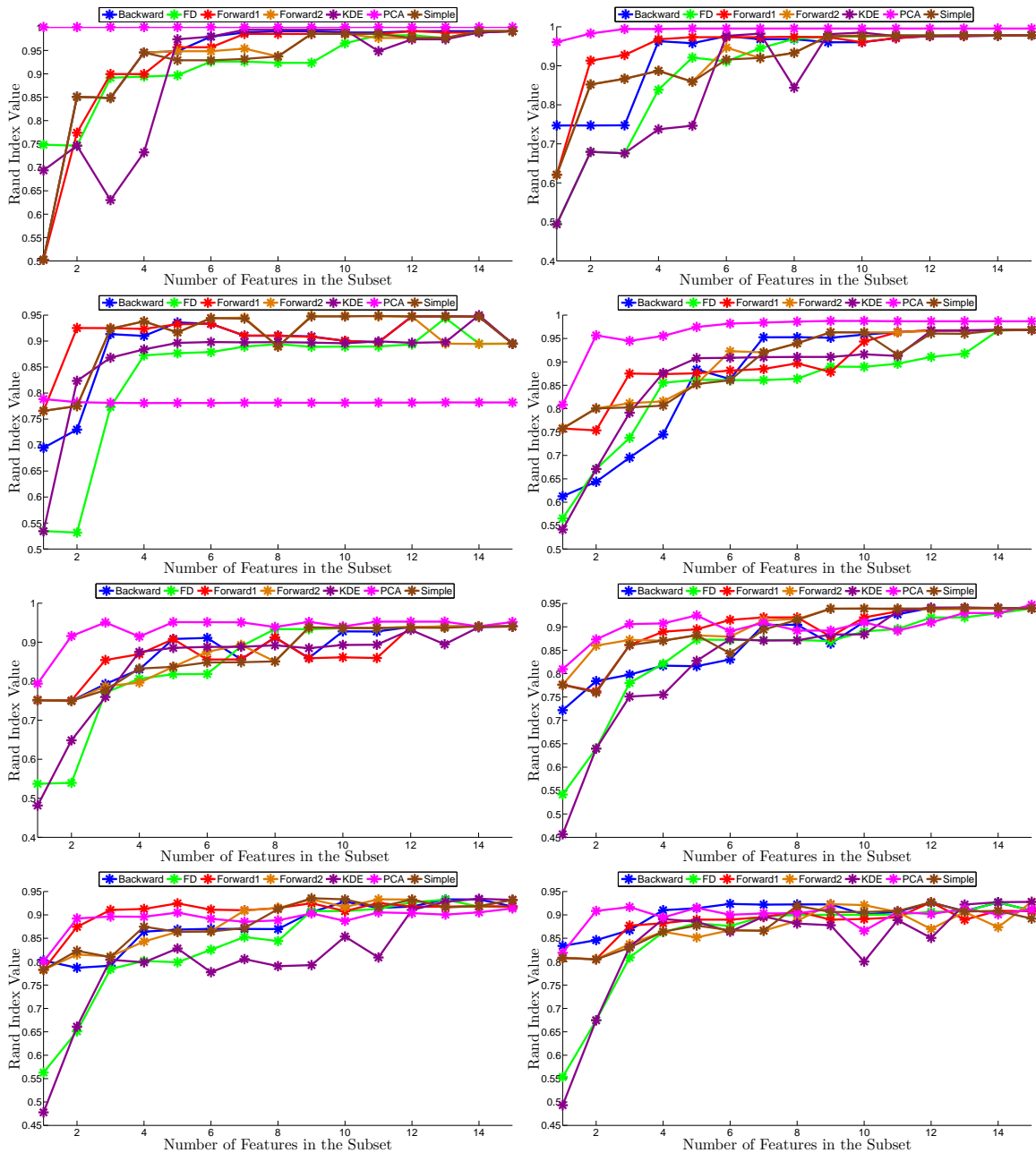


Figure B.5: Representation of the rand index for an increasing number of features, for each algorithm and for a noise level equal to 0.05. The maximum number of features is equal to 15 for each algorithm except PCA since the database does not count more than 15 features. For PCA, this stop level has been arbitrarily set to 15 too. The concordance is the following: Backward = Backward Elimination; Forward1 = Forward Selection 1; Forward2 = Forward Selection 2; Simple = Simple Ranking (for them, see subsection 4.4.1; KDE = Kernel Density Estimation of the Feature Density Similarity (see subsection 4.4.2); PCA = Principal Component Analysis. The disposition is the following: upper left = 2 clusters; upper right = 3 clusters; upper middle left = 4 clusters; upper middle right = 5 clusters; lower middle left = 6 clusters; lower middle right = 7 clusters; lower left = 8 clusters; lower right = 9 clusters.

## APPENDIX B. PERFORMANCES RESULTS

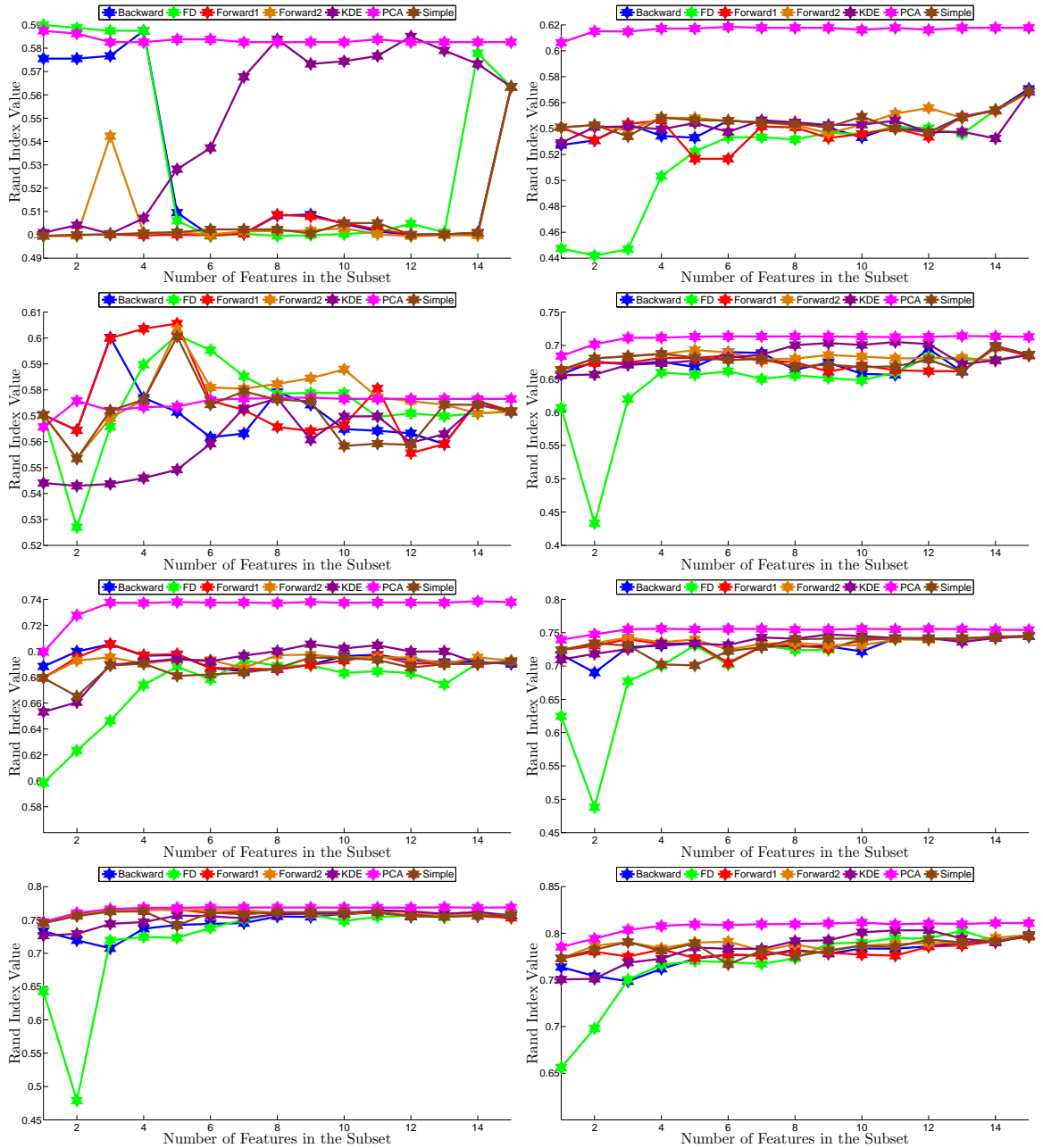


Figure B.6: Representation of the rand index for an increasing number of features, for each algorithm and for a noise level equal to 0.30. The maximum number of features is equal to 15 for each algorithm except PCA since the database does not count more than 15 features. For PCA, this stop level has been arbitrarily set to 15 too. The concordance is the following: Backward = Backward Elimination; Forward1 = Forward Selection 1; Forward2 = Forward Selection 2; Simple = Simple Ranking (for them, see subsection 4.4.1; KDE = Kernel Density Estimation of the Feature Density Similarity (see subsection 4.4.2); PCA = Principal Component Analysis. The disposition is the following: upper left = 2 clusters; upper right = 3 clusters; upper middle left = 4 clusters; upper middle right = 5 clusters; lower middle left = 6 clusters; lower middle right = 7 clusters; lower left = 8 clusters; lower right = 9 clusters.



## SELECTION OF THE RELEVANT FEATURES

This appendix shows the graphs of performance levels for each algorithm with the associated graphs of  $L_{ratio}$  and  $R_{correlation}$  .

## APPENDIX C. SELECTION OF THE RELEVANT FEATURES

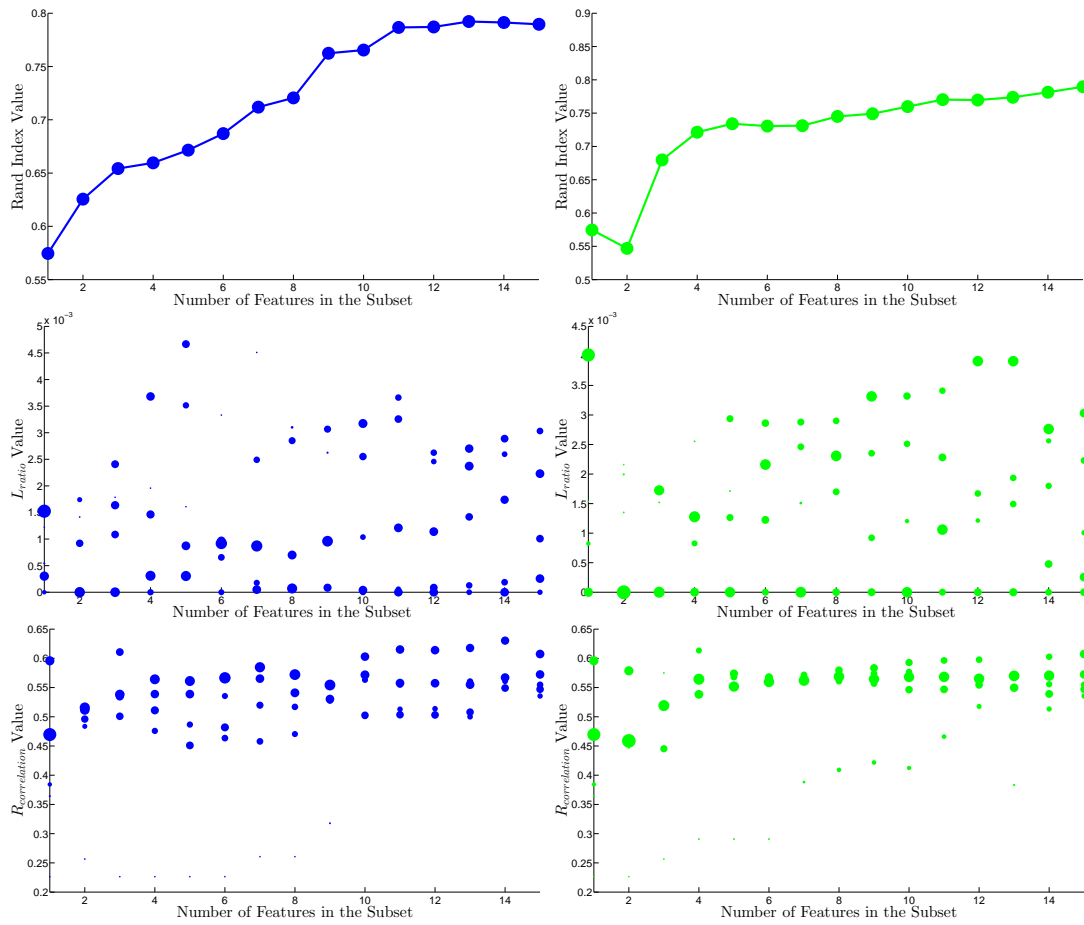


Figure C.1: Representation of Rand Index (upper graph),  $L_{ratio}$  (graph in the middle) and  $R$  (lower graph) values according to the number of features in the feature subset for Backward Elimination (left in blue) and Feature Dispersion (right in green), for 5 clusters and a 0.15 noise level.

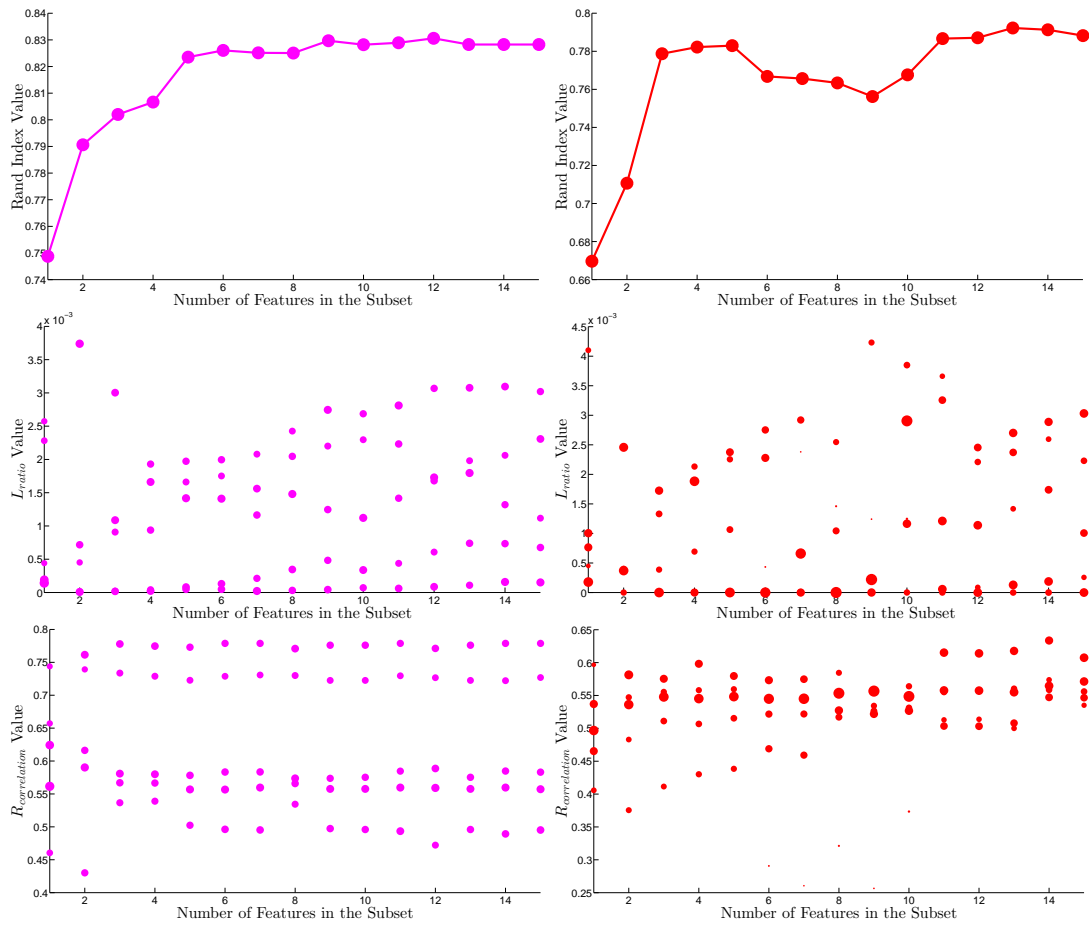


Figure C.2: Representation of Rand Index (upper graph),  $L_{ratio}$  (graph in the middle) and  $R$  (lower graph) values according to the number of features in the feature subset for Principal Component Analysis (left in pink) and Forward Selection 1 (right in red), for 5 clusters and a 0.15 noise level.

## APPENDIX C. SELECTION OF THE RELEVANT FEATURES

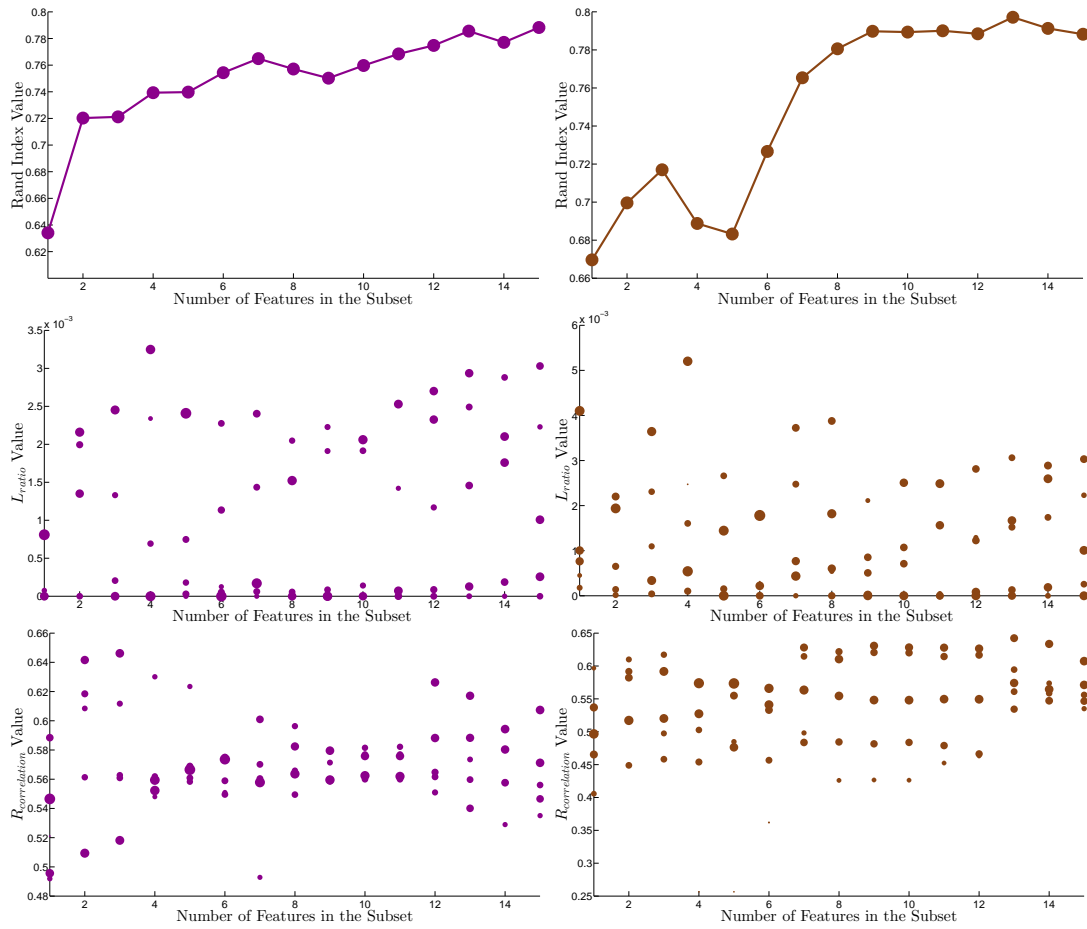


Figure C.3: Representation of Rand Index (upper graph),  $L_{ratio}$  (graph in the middle) and  $R$  (lower graph) values according to the number of features in the feature subset for Kernel Density Estimation (left in dark magenta) and Simple Ranking (right in brown), for 5 clusters and a 0.15 noise level.