

TABLE I. THE TABLE SHOWS THE CLASSIFICATION OF SUBJECTS AT THE CUT POINT 5% ON TEST-METHOD LEVEL USING JACCARD DISTANCE. THE FOUR VALUES IN EACH CELL ARE THE NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY TWO TECHNIQUES ARE HIGHLY DISSIMILAR, DISSIMILAR, SIMILAR AND HIGHLY SIMILAR RESPECTIVELY. THE LAST ROW SHOWS FOR EACH TECHNIQUE, THE TOTAL NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY THIS TECHNIQUE AND OTHER TECHNIQUES ARE HIGHLY ORTHOGONAL, ORTHOGONAL, SIMILAR, AND HIGHLY SIMILAR RESPECTIVELY.

Techniques	TP _{cg}	TP _{cg-add}	TP _{str}	TP _{topic-r}	TP _{topic-m}	TP _{total}	TP _{add}	TP _{art}	TP _{search}
TP _{cg}	- - - -	3 8 14 5	10 12 4 4	10 13 5 2	12 11 5 2	5 6 6 13	7 9 9 5	14 11 3 2	5 11 8 6
TP _{cg-add}	3 8 14 5	- - - -	7 14 6 3	7 14 7 2	5 17 8 0	8 8 9 5	3 13 10 4	10 13 6 1	3 12 9 6
TP _{str}	10 12 4 4	7 14 6 3	- - - -	6 11 10 3	2 6 6 16	8 13 5 4	7 11 10 2	12 13 4 1	6 12 10 2
TP _{topic-r}	10 13 5 2	7 14 7 2	6 11 10 3	- - - -	3 16 9 2	7 12 10 1	5 18 6 1	8 14 5 3	5 16 8 1
TP _{topic-m}	12 11 5 2	5 17 8 0	2 6 6 16	3 16 9 2	- - - -	6 13 7 4	3 11 13 3	10 15 5 0	4 15 11 0
TP _{total}	5 6 6 13	8 8 9 5	8 13 5 4	7 12 10 1	6 13 7 4	- - - -	3 9 8 10	16 9 4 1	4 10 9 7
TP _{add}	7 9 9 5	3 13 10 4	7 11 10 2	5 18 6 1	3 11 13 3	3 9 8 10	- - - -	12 9 8 1	1 4 8 17
TP _{art}	14 11 3 2	10 13 6 1	12 13 4 1	8 14 5 3	10 15 5 0	16 9 4 1	12 9 8 1	- - - -	14 10 5 1
TP _{search}	5 11 8 6	3 12 9 6	6 12 10 2	5 16 8 1	4 15 11 0	4 10 9 7	1 4 8 17	14 10 5 1	- - - -
Total	66 81 54 39	46 99 69 26	58 92 55 35	51 114 60 15	45 104 64 27	57 80 58 45	41 84 72 43	96 94 40 10	42 90 68 40

TABLE II. THE TABLE SHOWS THE CLASSIFICATION OF SUBJECTS AT THE CUT POINT 10% ON TEST-METHOD LEVEL USING JACCARD DISTANCE. THE FOUR VALUES IN EACH CELL ARE THE NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY TWO TECHNIQUES ARE HIGHLY DISSIMILAR, DISSIMILAR, SIMILAR AND HIGHLY SIMILAR RESPECTIVELY. THE LAST ROW SHOWS FOR EACH TECHNIQUE, THE TOTAL NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY THIS TECHNIQUE AND OTHER TECHNIQUES ARE HIGHLY ORTHOGONAL, ORTHOGONAL, SIMILAR, AND HIGHLY SIMILAR RESPECTIVELY.

Techniques	TP _{cg}	TP _{cg-add}	TP _{str}	TP _{topic-r}	TP _{topic-m}	TP _{total}	TP _{add}	TP _{art}	TP _{search}
TP _{cg}	- - - -	2 8 16 4	6 12 7 5	5 15 8 2	7 11 7 5	3 8 9 10	3 9 12 6	10 13 6 1	1 12 11 6
TP _{cg-add}	2 8 16 4	- - - -	4 11 10 5	2 12 12 4	3 12 13 2	1 12 11 6	1 9 11 9	7 12 9 2	2 9 11 8
TP _{str}	6 12 7 5	4 11 10 5	- - - -	3 13 13 1	0 2 10 18	5 8 12 5	2 9 10 9	9 12 9 0	3 8 12 7
TP _{topic-r}	5 15 8 2	2 12 12 4	3 13 13 1	- - - -	3 14 12 1	4 15 8 3	3 14 10 3	4 17 7 2	2 14 10 4
TP _{topic-m}	7 11 7 5	3 12 13 2	0 2 10 18	3 14 12 1	- - - -	4 11 11 4	3 3 19 5	5 14 11 0	4 4 18 4
TP _{total}	3 8 9 10	1 12 11 6	5 8 12 5	4 15 8 3	4 11 11 4	- - - -	1 9 11 9	11 14 4 1	2 9 10 9
TP _{add}	3 9 12 6	1 9 11 9	2 9 10 9	3 14 10 3	3 3 19 5	1 9 11 9	- - - -	5 12 11 2	1 2 5 22
TP _{art}	10 13 6 1	7 12 9 2	9 12 9 0	4 17 7 2	5 14 11 0	11 14 4 1	5 12 11 2	- - - -	6 12 10 2
TP _{search}	1 12 11 6	2 9 11 8	3 8 12 7	2 14 10 4	4 4 18 4	2 9 10 9	1 2 5 22	6 12 10 2	- - - -
Total	37 88 76 39	22 85 93 40	32 75 83 50	26 114 80 20	29 71 101 39	31 86 76 47	19 67 89 65	57 106 67 10	21 70 87 62

TABLE III. THE TABLE SHOWS THE CLASSIFICATION OF SUBJECTS AT THE CUT POINT 15% ON TEST-METHOD LEVEL USING JACCARD DISTANCE. THE FOUR VALUES IN EACH CELL ARE THE NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY TWO TECHNIQUES ARE HIGHLY DISSIMILAR, DISSIMILAR, SIMILAR AND HIGHLY SIMILAR RESPECTIVELY. THE LAST ROW SHOWS FOR EACH TECHNIQUE, THE TOTAL NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY THIS TECHNIQUE AND OTHER TECHNIQUES ARE HIGHLY ORTHOGONAL, ORTHOGONAL, SIMILAR, AND HIGHLY SIMILAR RESPECTIVELY.

Techniques	TP _{cg}	TP _{cg-add}	TP _{str}	TP _{topic-r}	TP _{topic-m}	TP _{total}	TP _{add}	TP _{art}	TP _{search}
TP _{cg}	- - - -	2 7 15 6	6 9 9 6	4 11 12 3	6 10 8 6	3 8 9 10	2 7 13 8	6 11 12 1	1 10 12 7
TP _{cg-add}	2 7 15 6	- - - -	1 11 14 4	1 10 16 3	1 9 14 6	2 10 11 7	1 7 12 10	4 10 13 3	2 8 11 9
TP _{str}	6 9 9 6	1 11 14 4	- - - -	2 10 11 7	0 2 8 20	2 12 11 5	2 5 13 10	3 14 13 0	3 3 17 7
TP _{topic-r}	4 11 12 3	1 10 16 3	2 10 11 7	- - - -	2 9 12 7	3 8 15 4	1 7 18 4	3 11 13 3	2 7 18 3
TP _{topic-m}	6 10 8 6	1 9 14 6	0 2 8 20	2 9 12 7	- - - -	3 8 14 5	3 2 18 7	2 13 15 0	4 1 17 8
TP _{total}	3 8 9 10	2 10 11 7	2 12 11 5	3 8 15 4	3 8 14 5	- - - -	1 7 11 11	9 8 12 1	2 6 11 11
TP _{add}	2 7 13 8	1 7 12 10	2 5 13 10	1 7 18 4	3 2 18 7	1 7 11 11	- - - -	2 10 13 5	0 1 5 24
TP _{art}	6 11 12 1	4 10 13 3	3 14 13 0	3 11 13 3	2 13 15 0	9 8 12 1	2 10 13 5	- - - -	2 11 12 5
TP _{search}	1 10 12 7	2 8 11 9	3 3 17 7	2 7 18 3	4 1 17 8	2 6 11 11	0 1 5 24	2 11 12 5	- - - -
Total	30 73 90 47	14 72 106 48	19 66 96 59	18 73 115 34	21 54 106 59	25 67 94 54	12 46 103 79	31 88 103 18	16 47 103 74

TABLE IV. THE TABLE SHOWS THE CLASSIFICATION OF SUBJECTS AT THE CUT POINT 20% ON TEST-METHOD LEVEL USING JACCARD DISTANCE. THE FOUR VALUES IN EACH CELL ARE THE NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY TWO TECHNIQUES ARE HIGHLY DISSIMILAR, DISSIMILAR, SIMILAR AND HIGHLY SIMILAR RESPECTIVELY. THE LAST ROW SHOWS FOR EACH TECHNIQUE, THE TOTAL NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY THIS TECHNIQUE AND OTHER TECHNIQUES ARE HIGHLY ORTHOGONAL, ORTHOGONAL, SIMILAR, AND HIGHLY SIMILAR RESPECTIVELY.

Techniques	TP _{cg}	TP _{cg-add}	TP _{str}	TP _{topic-r}	TP _{topic-m}	TP _{total}	TP _{add}	TP _{art}	TP _{search}
TP _{cg}	- - - -	2 4 17 7	5 6 13 6	2 8 15 5	3 5 16 6	2 3 15 10	1 6 15 8	2 11 14 3	1 7 14 8
TP _{cg-add}	2 4 17 7	- - - -	1 8 13 8	1 6 17 6	1 5 13 11	3 9 9 9	1 5 12 12	3 5 16 6	1 4 12 13
TP _{str}	5 6 13 6	1 8 13 8	- - - -	2 7 12 9	1 0 4 25	2 8 13 7	2 1 14 13	4 7 16 3	1 1 15 13
TP _{topic-r}	2 8 15 5	1 6 17 6	2 7 12 9	- - - -	1 6 10 13	3 11 10 6	0 7 18 5	1 8 17 4	0 7 18 5
TP _{topic-m}	3 5 16 6	1 5 13 11	1 0 4 25	1 6 10 13	- - - -	1 7 15 7	1 1 12 16	1 8 18 3	1 0 12 17
TP _{total}	2 3 15 10	3 9 9 9	2 8 13 7	3 11 10 6	1 7 15 7	- - - -	1 8 8 13	4 10 13 3	0 7 9 14
TP _{add}	1 6 15 8	1 5 12 12	2 1 14 13	0 7 18 5	1 1 12 16	1 8 8 13	- - - -	0 7 14 9	0 0 3 27
TP _{art}	2 11 14 3	3 5 16 6	4 7 16 3	1 8 17 4	1 8 18 3	4 10 13 3	0 7 14 9	- - - -	0 6 16 8
TP _{search}	1 7 14 8	1 4 12 13	1 1 15 13	0 7 18 5	1 0 12 17	0 7 9 14	0 0 3 27	0 6 16 8	- - - -
Total	18 50 119 53	13 46 109 72	18 38 100 84	10 60 117 53	10 32 100 98	16 63 92 69	6 35 96 103	15 62 124 39	4 32 99 105

TABLE V. THE TABLE SHOWS THE CLASSIFICATION OF SUBJECTS AT THE CUT POINT 25% ON TEST-METHOD LEVEL USING JACCARD DISTANCE. THE FOUR VALUES IN EACH CELL ARE THE NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY TWO TECHNIQUES ARE HIGHLY DISSIMILAR, DISSIMILAR, SIMILAR AND HIGHLY SIMILAR RESPECTIVELY. THE LAST ROW SHOWS FOR EACH TECHNIQUE, THE TOTAL NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY THIS TECHNIQUE AND OTHER TECHNIQUES ARE HIGHLY ORTHOGONAL, ORTHOGONAL, SIMILAR, AND HIGHLY SIMILAR RESPECTIVELY.

Techniques	TP _{cg}	TP _{cg-add}	TP _{str}	TP _{topic-r}	TP _{topic-m}	TP _{total}	TP _{add}	TP _{art}	TP _{search}
TP _{cg}	- - - -	2 5 15 8	2 5 14 9	3 6 17 4	2 5 17 6	1 3 13 13	1 5 16 8	2 7 19 2	1 7 12 10
TP _{cg-add}	2 5 15 8	- - - -	0 7 13 10	0 8 14 8	1 4 12 13	1 8 12 9	0 4 11 15	1 5 18 6	0 4 11 15
TP _{str}	2 5 14 9	0 7 13 10	- - - -	1 6 13 10	0 0 3 27	1 3 15 11	0 1 11 18	2 3 21 4	0 1 12 17
TP _{topic-r}	3 6 17 4	0 8 14 8	1 6 13 10	- - - -	1 4 12 13	1 13 11 5	0 5 18 7	0 4 20 6	0 3 17 10
TP _{topic-m}	2 5 17 6	1 4 12 13	0 0 3 27	1 4 12 13	- - - -	1 3 18 8	1 1 8 20	1 6 17 6	1 0 11 18
TP _{total}	1 3 13 13	1 8 12 9	1 3 15 11	1 13 11 5	1 3 18 8	- - - -	0 4 10 16	3 6 17 4	0 5 8 17
TP _{add}	1 5 16 8	0 4 11 15	0 1 11 18	0 5 18 7	1 1 8 20	0 4 10 16	- - - -	1 1 15 13	0 0 3 27
TP _{art}	2 7 19 2	1 5 18 6	2 3 21 4	0 4 20 6	1 6 17 6	3 6 17 4	1 1 15 13	- - - -	1 1 17 11
TP _{search}	1 7 12 10	0 4 11 15	0 1 12 17	0 3 17 10	1 0 11 18	0 5 8 17	0 0 3 27	1 1 17 11	- - - -
Total	14 43 123 60	5 45 106 84	6 26 102 106	6 49 122 63	8 23 98 111	8 45 104 83	3 21 92 124	11 33 144 52	3 21 91 125

TABLE VI. THE TABLE SHOWS THE CLASSIFICATION OF SUBJECTS AT THE CUT POINT 30% ON TEST-METHOD LEVEL USING JACCARD DISTANCE. THE FOUR VALUES IN EACH CELL ARE THE NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY TWO TECHNIQUES ARE HIGHLY DISSIMILAR, DISSIMILAR, SIMILAR AND HIGHLY SIMILAR RESPECTIVELY. THE LAST ROW SHOWS FOR EACH TECHNIQUE, THE TOTAL NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY THIS TECHNIQUE AND OTHER TECHNIQUES ARE HIGHLY ORTHOGONAL, ORTHOGONAL, SIMILAR, AND HIGHLY SIMILAR RESPECTIVELY.

Techniques	TP _{cg}	TP _{cg-add}	TP _{str}	TP _{topic-r}	TP _{topic-m}	TP _{total}	TP _{add}	TP _{art}	TP _{search}
TP _{cg}	- - - -	0 6 16 8	2 4 16 8	2 6 16 6	2 3 18 7	0 4 13 13	0 5 15 10	2 6 17 5	0 5 14 11
TP _{cg-add}	0 6 16 8	- - - -	0 3 14 13	0 4 17 9	0 2 14 14	1 5 16 8	0 3 8 19	1 4 17 8	0 3 11 16
TP _{str}	2 4 16 8	0 3 14 13	- - - -	1 4 13 12	0 0 1 29	1 3 13 13	0 1 9 20	2 3 16 9	0 1 10 19
TP _{topic-r}	2 6 16 6	0 4 17 9	1 4 13 12	- - - -	1 4 12 13	0 9 14 7	0 3 17 10	0 3 18 9	0 2 17 11
TP _{topic-m}	2 3 18 7	0 2 14 14	0 0 1 29	1 4 12 13	- - - -	1 3 18 8	0 1 6 23	1 3 16 10	0 1 7 22
TP _{total}	0 4 13 13	1 5 16 8	1 3 13 13	0 9 14 7	1 3 18 8	- - - -	0 2 11 17	2 4 21 3	0 3 9 18
TP _{add}	0 5 15 10	0 3 8 19	0 1 9 20	0 3 17 10	0 1 6 23	0 2 11 17	- - - -	1 1 14 14	0 0 1 29
TP _{art}	2 6 17 5	1 4 17 8	2 3 16 9	0 3 18 9	1 3 16 10	2 4 21 3	1 1 14 14	- - - -	1 1 14 14
TP _{search}	0 5 14 11	0 3 11 16	0 1 10 19	0 2 17 11	0 1 7 22	0 3 9 18	0 0 1 29	1 1 14 14	- - - -
Total	8 39 125 68	2 30 113 95	6 19 92 123	4 35 124 77	5 17 92 126	5 33 115 87	1 16 81 142	10 25 133 72	1 16 83 140

TABLE VII. THE TABLE SHOWS THE CLASSIFICATION OF SUBJECTS AT THE CUT POINT 35% ON TEST-METHOD LEVEL USING JACCARD DISTANCE. THE FOUR VALUES IN EACH CELL ARE THE NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY TWO TECHNIQUES ARE HIGHLY DISSIMILAR, DISSIMILAR, SIMILAR AND HIGHLY SIMILAR RESPECTIVELY. THE LAST ROW SHOWS FOR EACH TECHNIQUE, THE TOTAL NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY THIS TECHNIQUE AND OTHER TECHNIQUES ARE HIGHLY ORTHOGONAL, ORTHOGONAL, SIMILAR, AND HIGHLY SIMILAR RESPECTIVELY.

Techniques	TP _{cg}	TP _{cg-add}	TP _{str}	TP _{topic-r}	TP _{topic-m}	TP _{total}	TP _{add}	TP _{art}	TP _{search}
TP _{cg}	- - - -	0 5 15 10	2 3 17 8	1 6 15 8	2 3 17 8	0 4 8 18	0 5 14 11	2 4 18 6	0 5 13 12
TP _{cg-add}	0 5 15 10	- - - -	0 2 12 16	0 3 13 14	0 2 14 14	1 4 12 13	0 3 8 19	0 6 13 11	0 3 9 18
TP _{str}	2 3 17 8	0 2 12 16	- - - -	0 2 14 14	0 0 2 28	1 1 13 15	0 1 7 22	2 2 12 14	0 1 9 20
TP _{topic-r}	1 6 15 8	0 3 13 14	0 2 14 14	- - - -	0 2 13 15	0 5 14 11	0 0 14 16	0 2 11 17	0 0 16 14
TP _{topic-m}	2 3 17 8	0 2 14 14	0 0 2 28	0 2 13 15	- - - -	1 0 15 14	0 1 5 24	1 3 10 16	0 1 6 23
TP _{total}	0 4 8 18	1 4 12 13	1 1 13 15	0 5 14 11	1 0 15 14	- - - -	0 2 9 19	1 5 18 6	0 2 9 19
TP _{add}	0 5 14 11	0 3 8 19	0 1 7 22	0 0 14 16	0 1 5 24	0 2 9 19	- - - -	0 2 8 20	0 0 1 29
TP _{art}	2 4 18 6	0 6 13 11	2 2 12 14	0 2 11 17	1 3 10 16	1 5 18 6	0 2 8 20	- - - -	0 2 8 20
TP _{search}	0 5 13 12	0 3 9 18	0 1 9 20	0 0 16 14	0 1 6 23	0 2 9 19	0 0 1 29	0 2 8 20	- - - -
Total	7 35 117 81	1 28 96 115	5 12 86 137	1 20 110 109	4 12 82 142	4 23 98 115	0 14 66 160	6 26 98 110	0 14 71 155

TABLE VIII. THE TABLE SHOWS THE CLASSIFICATION OF SUBJECTS AT THE CUT POINT 40% ON TEST-METHOD LEVEL USING JACCARD DISTANCE. THE FOUR VALUES IN EACH CELL ARE THE NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY TWO TECHNIQUES ARE HIGHLY DISSIMILAR, DISSIMILAR, SIMILAR AND HIGHLY SIMILAR RESPECTIVELY. THE LAST ROW SHOWS FOR EACH TECHNIQUE, THE TOTAL NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY THIS TECHNIQUE AND OTHER TECHNIQUES ARE HIGHLY ORTHOGONAL, ORTHOGONAL, SIMILAR, AND HIGHLY SIMILAR RESPECTIVELY.

Techniques	TP _{cg}	TP _{cg-add}	TP _{str}	TP _{topic-r}	TP _{topic-m}	TP _{total}	TP _{add}	TP _{art}	TP _{search}
TP _{cg}	- - - -	0 2 14 14	0 3 16 11	1 2 16 11	1 2 13 14	0 2 7 21	0 1 16 13	1 2 19 8	0 1 15 14
TP _{cg-add}	0 2 14 14	- - - -	0 1 10 19	0 2 11 17	0 1 11 18	1 1 12 16	0 1 5 24	0 2 10 18	0 2 4 24
TP _{str}	0 3 16 11	0 1 10 19	- - - -	0 1 12 17	0 1 1 28	0 2 9 19	0 0 5 25	0 2 10 18	0 0 5 25
TP _{topic-r}	1 2 16 11	0 2 11 17	0 1 12 17	- - - -	0 0 10 20	0 4 12 14	0 0 9 21	0 1 8 21	0 0 9 21
TP _{topic-m}	1 2 13 14	0 1 11 18	0 1 1 28	0 0 10 20	- - - -	0 1 10 19	0 0 6 24	0 1 7 22	0 0 6 24
TP _{total}	0 2 7 21	1 1 12 16	0 2 9 19	0 4 12 14	0 1 10 19	- - - -	0 0 10 20	0 3 16 11	0 0 10 20
TP _{add}	0 1 16 13	0 1 5 24	0 0 5 25	0 0 9 21	0 0 6 24	0 0 10 20	- - - -	0 1 4 25	0 0 1 29
TP _{art}	1 2 19 8	0 2 10 18	0 2 10 18	0 1 8 21	0 1 7 22	0 3 16 11	0 1 4 25	- - - -	0 1 3 26
TP _{search}	0 1 15 14	0 2 4 24	0 0 5 25	0 0 9 21	0 0 6 24	0 0 10 20	0 0 1 29	0 1 3 26	- - - -
Total	3 15 116 106	1 12 77 150	0 10 68 162	1 10 87 142	1 6 64 169	1 13 86 140	0 3 56 181	1 13 77 149	0 4 53 183

TABLE IX. THE TABLE SHOWS THE CLASSIFICATION OF SUBJECTS AT THE CUT POINT 45% ON TEST-METHOD LEVEL USING JACCARD DISTANCE. THE FOUR VALUES IN EACH CELL ARE THE NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY TWO TECHNIQUES ARE HIGHLY DISSIMILAR, DISSIMILAR, SIMILAR AND HIGHLY SIMILAR RESPECTIVELY. THE LAST ROW SHOWS FOR EACH TECHNIQUE, THE TOTAL NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY THIS TECHNIQUE AND OTHER TECHNIQUES ARE HIGHLY ORTHOGONAL, ORTHOGONAL, SIMILAR, AND HIGHLY SIMILAR RESPECTIVELY.

Techniques	TP _{cg}	TP _{cg-add}	TP _{str}	TP _{topic-r}	TP _{topic-m}	TP _{total}	TP _{add}	TP _{art}	TP _{search}
TP _{cg}	- - - -	1 0 11 18	0 1 13 16	1 1 17 11	1 0 14 15	0 1 5 24	0 1 10 19	1 1 17 11	0 1 11 18
TP _{cg-add}	1 0 11 18	- - - -	0 1 8 21	0 2 11 17	0 1 7 22	1 0 11 18	0 1 5 24	0 1 8 21	0 1 5 24
TP _{str}	0 1 13 16	0 1 8 21	- - - -	0 1 11 18	0 1 1 28	0 1 10 19	0 0 5 25	0 1 6 23	0 0 4 26
TP _{topic-r}	1 1 17 11	0 2 11 17	0 1 11 18	- - - -	0 0 9 21	0 2 13 15	0 0 10 20	0 1 9 20	0 0 10 20
TP _{topic-m}	1 0 14 15	0 1 7 22	0 1 1 28	0 0 9 21	- - - -	0 1 11 18	0 0 3 27	0 0 5 25	0 0 3 27
TP _{total}	0 1 5 24	1 0 11 18	0 1 10 19	0 2 13 15	0 1 11 18	- - - -	0 0 8 22	0 1 15 14	0 0 8 22
TP _{add}	0 1 10 19	0 1 5 24	0 0 5 25	0 0 10 20	0 0 3 27	0 0 8 22	- - - -	0 0 3 27	0 0 1 29
TP _{art}	1 1 17 11	0 1 8 21	0 1 6 23	0 1 9 20	0 0 5 25	0 1 15 14	0 0 3 27	- - - -	0 0 3 27
TP _{search}	0 1 11 18	0 1 5 24	0 0 4 26	0 0 10 20	0 0 3 27	0 0 8 22	0 0 1 29	0 0 3 27	- - - -
Total	4 6 98 132	2 7 66 165	0 6 58 176	1 7 90 142	1 3 53 183	1 6 81 152	0 2 45 193	1 5 66 168	0 2 45 193

TABLE X. THE TABLE SHOWS THE CLASSIFICATION OF SUBJECTS AT THE CUT POINT 50% ON TEST-METHOD LEVEL USING JACCARD DISTANCE. THE FOUR VALUES IN EACH CELL ARE THE NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY TWO TECHNIQUES ARE HIGHLY DISSIMILAR, DISSIMILAR, SIMILAR AND HIGHLY SIMILAR RESPECTIVELY. THE LAST ROW SHOWS FOR EACH TECHNIQUE, THE TOTAL NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY THIS TECHNIQUE AND OTHER TECHNIQUES ARE HIGHLY ORTHOGONAL, ORTHOGONAL, SIMILAR, AND HIGHLY SIMILAR RESPECTIVELY.

Techniques	TP _{cg}	TP _{cg-add}	TP _{str}	TP _{topic-r}	TP _{topic-m}	TP _{total}	TP _{add}	TP _{art}	TP _{search}
TP _{cg}	- - - -	1 0 6 23	0 1 8 21	1 1 15 13	1 1 7 21	0 1 3 26	0 1 7 22	1 0 10 19	0 1 7 22
TP _{cg-add}	1 0 6 23	- - - -	0 1 5 24	0 1 8 21	0 0 8 22	0 1 6 23	0 0 4 26	0 0 5 25	0 0 4 26
TP _{str}	0 1 8 21	0 1 5 24	- - - -	0 0 10 20	0 0 2 28	0 0 7 23	0 0 2 28	0 0 5 25	0 0 3 27
TP _{topic-r}	1 1 15 13	0 1 8 21	0 0 10 20	- - - -	0 0 6 24	0 2 12 16	0 0 7 23	0 0 8 22	0 0 8 22
TP _{topic-m}	1 1 7 21	0 0 8 22	0 0 2 28	0 0 6 24	- - - -	0 0 10 20	0 0 3 27	0 0 2 28	0 0 3 27
TP _{total}	0 1 3 26	0 1 6 23	0 0 7 23	0 2 12 16	0 0 10 20	- - - -	0 0 7 23	0 0 13 17	0 0 6 24
TP _{add}	0 1 7 22	0 0 4 26	0 0 2 28	0 0 7 23	0 0 3 27	0 0 7 23	- - - -	0 0 3 27	0 0 1 29
TP _{art}	1 0 10 19	0 0 5 25	0 0 5 25	0 0 8 22	0 0 2 28	0 0 13 17	0 0 3 27	- - - -	0 0 1 29
TP _{search}	0 1 7 22	0 0 4 26	0 0 3 27	0 0 8 22	0 0 3 27	0 0 6 24	0 0 1 29	0 0 1 29	- - - -
Total	4 6 63 167	1 3 46 190	0 2 42 196	1 4 74 161	1 1 41 197	0 4 64 172	0 1 34 205	1 0 47 192	0 1 33 206

TABLE XI. THE TABLE SHOWS THE CLASSIFICATION OF SUBJECTS AT THE CUT POINT 5% ON TEST-CLASS LEVEL USING JACCARD DISTANCE. THE FOUR VALUES IN EACH CELL ARE THE NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY TWO TECHNIQUES ARE HIGHLY DISSIMILAR, DISSIMILAR, SIMILAR AND HIGHLY SIMILAR RESPECTIVELY. THE LAST ROW SHOWS FOR EACH TECHNIQUE, THE TOTAL NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY THIS TECHNIQUE AND OTHER TECHNIQUES ARE HIGHLY ORTHOGONAL, ORTHOGONAL, SIMILAR, AND HIGHLY SIMILAR RESPECTIVELY.

Techniques	TP _{cg}	TP _{cg-add}	TP _{str}	TP _{topic-r}	TP _{topic-m}	TP _{total}	TP _{add}	TP _{art}	TP _{search}
TP _{cg}	- - -	0 2 5 23	15 5 7 3	20 5 1 4	16 8 1 5	8 9 7 6	9 9 4 8	20 5 1 4	9 10 3 8
TP _{cg-add}	0 2 5 23	- - -	14 9 5 2	20 5 1 4	15 9 1 5	10 8 5 7	11 8 4 7	19 7 0 4	10 9 4 7
TP _{str}	15 5 7 3	14 9 5 2	- - -	15 6 4 5	10 5 7 8	14 9 4 3	16 5 4 5	19 4 1 6	16 5 4 5
TP _{topic-r}	20 5 1 4	20 5 1 4	15 6 4 5	- - -	7 1 3 19	18 8 2 2	18 7 3 2	19 5 4 2	18 8 2 2
TP _{topic-m}	16 8 1 5	15 9 1 5	10 5 7 8	7 1 3 19	- - -	15 10 2 3	16 9 2 3	17 9 2 2	17 8 2 3
TP _{total}	8 9 7 6	10 8 5 7	14 9 4 3	18 8 2 2	15 10 2 3	- - -	0 2 4 24	18 8 2 2	0 2 5 23
TP _{add}	9 9 4 8	11 8 4 7	16 5 4 5	18 7 3 2	16 9 2 3	0 2 4 24	- - -	19 6 2 3	0 0 2 28
TP _{art}	20 5 1 4	19 7 0 4	19 4 1 6	19 5 4 2	17 9 2 2	18 8 2 2	19 6 2 3	- - -	19 6 2 3
TP _{search}	9 10 3 8	10 9 4 7	16 5 4 5	18 8 2 2	17 8 2 3	0 2 5 23	0 0 2 28	19 6 2 3	- - -
Total	97 53 29 61	99 57 25 59	119 48 36 37	135 45 20 40	113 59 20 48	83 56 31 70	89 46 25 80	150 50 14 26	89 48 24 79

TABLE XII. THE TABLE SHOWS THE CLASSIFICATION OF SUBJECTS AT THE CUT POINT 10% ON TEST-CLASS LEVEL USING JACCARD DISTANCE. THE FOUR VALUES IN EACH CELL ARE THE NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY TWO TECHNIQUES ARE HIGHLY DISSIMILAR, DISSIMILAR, SIMILAR AND HIGHLY SIMILAR RESPECTIVELY. THE LAST ROW SHOWS FOR EACH TECHNIQUE, THE TOTAL NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY THIS TECHNIQUE AND OTHER TECHNIQUES ARE HIGHLY ORTHOGONAL, ORTHOGONAL, SIMILAR, AND HIGHLY SIMILAR RESPECTIVELY.

Techniques	TP _{cg}	TP _{cg-add}	TP _{str}	TP _{topic-r}	TP _{topic-m}	TP _{total}	TP _{add}	TP _{art}	TP _{search}
TP _{cg}	- - -	1 2 11 16	6 7 7 10	18 8 1 3	9 6 8 7	7 9 4 10	7 7 7 9	19 7 3 1	7 8 6 9
TP _{cg-add}	1 2 11 16	- - -	6 7 7 10	17 10 2 1	8 12 5 5	6 10 8 6	7 8 7 8	18 10 2 0	7 8 6 9
TP _{str}	6 7 7 10	6 7 7 10	- - -	16 6 6 2	5 7 7 11	10 6 8 6	8 7 9 6	17 7 2 4	8 7 8 7
TP _{topic-r}	18 8 1 3	17 10 2 1	16 6 6 2	- - -	9 7 7 7	16 7 4 3	17 6 5 2	17 6 4 3	17 7 3 3
TP _{topic-m}	9 6 8 7	8 12 5 5	5 7 7 11	9 7 7 7	- - -	10 11 3 6	10 10 4 6	16 8 4 2	10 10 4 6
TP _{total}	7 9 4 10	6 10 8 6	10 6 8 6	16 7 4 3	10 11 3 6	- - -	0 5 7 18	16 8 3 3	0 5 8 17
TP _{add}	7 7 7 9	7 8 7 8	8 7 9 6	17 6 5 2	10 10 4 6	0 5 7 18	- - -	13 9 5 3	0 0 1 29
TP _{art}	19 7 3 1	18 10 2 0	17 7 2 4	17 6 4 3	16 8 4 2	16 8 3 3	13 9 5 3	- - -	13 8 6 3
TP _{search}	7 8 6 9	7 8 6 9	8 7 8 7	17 7 3 3	10 10 4 6	0 5 8 17	0 0 1 29	13 8 6 3	- - -
Total	74 54 47 65	70 67 48 55	76 54 54 56	127 57 32 24	77 71 42 50	65 61 45 69	62 52 45 81	129 63 29 19	62 53 42 83

TABLE XIII. THE TABLE SHOWS THE CLASSIFICATION OF SUBJECTS AT THE CUT POINT 15% ON TEST-CLASS LEVEL USING JACCARD DISTANCE. THE FOUR VALUES IN EACH CELL ARE THE NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY TWO TECHNIQUES ARE HIGHLY DISSIMILAR, DISSIMILAR, SIMILAR AND HIGHLY SIMILAR RESPECTIVELY. THE LAST ROW SHOWS FOR EACH TECHNIQUE, THE TOTAL NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY THIS TECHNIQUE AND OTHER TECHNIQUES ARE HIGHLY ORTHOGONAL, ORTHOGONAL, SIMILAR, AND HIGHLY SIMILAR RESPECTIVELY.

Techniques	TP _{cg}	TP _{cg-add}	TP _{str}	TP _{topic-r}	TP _{topic-m}	TP _{total}	TP _{add}	TP _{art}	TP _{search}
TP _{cg}	- - -	0 4 8 18	4 4 9 13	13 8 3 6	7 5 11 7	4 10 5 11	5 7 9 9	13 11 6 0	5 7 9 9
TP _{cg-add}	0 4 8 18	- - -	4 6 7 13	13 6 7 4	7 9 8 6	4 8 9 9	5 6 9 10	13 10 4 3	4 7 8 11
TP _{str}	4 4 9 13	4 6 7 13	- - -	14 4 7 5	5 5 7 13	6 9 8 7	6 8 7 9	11 10 6 3	5 9 9 7
TP _{topic-r}	13 8 3 6	13 6 7 4	14 4 7 5	- - -	9 6 9 6	11 11 5 3	10 10 6 4	15 10 4 1	10 9 7 4
TP _{topic-m}	7 5 11 7	7 9 8 6	5 5 7 13	9 6 9 6	- - -	9 7 8 6	7 7 8 8	11 11 6 2	7 7 8 8
TP _{total}	4 10 5 11	4 8 9 9	6 9 8 7	11 11 5 3	9 7 8 6	- - -	0 3 10 17	13 12 4 1	0 3 10 17
TP _{add}	5 7 9 9	5 6 9 10	6 8 7 9	10 10 6 4	7 7 8 8	0 3 10 17	- - -	12 9 6 3	0 0 2 28
TP _{art}	13 11 6 0	13 10 4 3	11 10 6 3	15 10 4 1	11 11 6 2	13 12 4 1	12 9 6 3	- - -	12 9 7 2
TP _{search}	5 7 9 9	4 7 8 11	5 9 9 7	10 9 7 4	7 7 8 8	0 3 10 17	0 0 2 28	12 9 7 2	- - -
Total	51 56 60 73	50 56 60 74	55 55 60 70	95 64 48 33	62 57 65 56	47 63 59 71	45 50 57 88	100 82 43 15	43 51 60 86

TABLE XIV. THE TABLE SHOWS THE CLASSIFICATION OF SUBJECTS AT THE CUT POINT 20% ON TEST-CLASS LEVEL USING JACCARD DISTANCE. THE FOUR VALUES IN EACH CELL ARE THE NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY TWO TECHNIQUES ARE HIGHLY DISSIMILAR, DISSIMILAR, SIMILAR AND HIGHLY SIMILAR RESPECTIVELY. THE LAST ROW SHOWS FOR EACH TECHNIQUE, THE TOTAL NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY THIS TECHNIQUE AND OTHER TECHNIQUES ARE HIGHLY ORTHOGONAL, ORTHOGONAL, SIMILAR, AND HIGHLY SIMILAR RESPECTIVELY.

Techniques	TP _{cg}	TP _{cg-add}	TP _{str}	TP _{topic-r}	TP _{topic-m}	TP _{total}	TP _{add}	TP _{art}	TP _{search}
TP _{cg}	- - -	0 4 5 21	2 3 12 13	6 12 5 7	2 8 12 8	2 8 8 12	3 8 10 9	10 11 7 2	3 8 10 9
TP _{cg-add}	0 4 5 21	- - -	2 4 13 11	6 9 9 6	2 7 12 9	3 6 10 11	2 7 12 9	11 9 7 3	2 7 12 9
TP _{str}	2 3 12 13	2 4 13 11	- - -	9 8 8 5	1 4 11 14	4 4 14 8	4 5 11 10	8 10 7 5	4 5 10 11
TP _{topic-r}	6 12 5 7	6 9 9 6	9 8 8 5	- - -	6 10 6 8	7 12 7 4	7 12 7 4	12 10 4 4	7 12 7 4
TP _{topic-m}	2 8 12 8	2 7 12 9	1 4 11 14	6 10 6 8	- - -	5 8 7 10	5 3 12 10	7 13 7 3	5 3 11 11
TP _{total}	2 8 8 12	3 6 10 11	4 4 14 8	7 12 7 4	5 8 7 10	- - -	0 2 7 21	9 13 7 1	0 2 7 21
TP _{add}	3 8 10 9	2 7 12 9	4 5 11 10	7 12 7 4	5 3 12 10	2 2 7 21	- - -	9 9 9 3	0 0 0 30
TP _{art}	10 11 7 2	11 9 7 3	8 10 7 5	12 10 4 4	7 13 7 3	9 13 7 1	9 9 9 3	- - -	9 9 9 3
TP _{search}	3 8 10 9	2 7 12 9	4 5 10 11	7 12 7 4	5 3 11 11	0 2 7 21	0 0 0 30	9 9 9 3	- - -
Total	28 62 69 81	28 53 80 79	34 43 86 77	60 85 53 42	33 56 78 73	30 55 67 88	30 46 68 96	75 84 57 24	30 46 66 98

TABLE XV. THE TABLE SHOWS THE CLASSIFICATION OF SUBJECTS AT THE CUT POINT 25% ON TEST-CLASS LEVEL USING JACCARD DISTANCE. THE FOUR VALUES IN EACH CELL ARE THE NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY TWO TECHNIQUES ARE HIGHLY DISSIMILAR, DISSIMILAR, SIMILAR AND HIGHLY SIMILAR RESPECTIVELY. THE LAST ROW SHOWS FOR EACH TECHNIQUE, THE TOTAL NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY THIS TECHNIQUE AND OTHER TECHNIQUES ARE HIGHLY ORTHOGONAL, ORTHOGONAL, SIMILAR, AND HIGHLY SIMILAR RESPECTIVELY.

Techniques	TP _{cg}	TP _{cg-add}	TP _{str}	TP _{topic-r}	TP _{topic-m}	TP _{total}	TP _{add}	TP _{art}	TP _{search}
TP _{cg}	- - -	0 3 4 23	2 2 12 14	5 11 7 7	2 8 11 9	2 6 11 11	3 4 14 9	5 11 11 3	3 3 15 9
TP _{cg-add}	0 3 4 23	- - -	2 2 17 9	3 13 8 6	2 7 10 11	3 6 12 9	2 8 10 10	6 8 12 4	2 8 10 10
TP _{str}	2 2 12 14	2 2 17 9	- - -	7 9 8 6	1 3 12 14	4 4 12 10	4 4 11 11	3 10 11 6	4 3 13 10
TP _{topic-r}	5 11 7 7	3 13 8 6	7 9 8 6	- - -	2 9 10 9	4 11 11 4	3 11 12 4	8 11 6 5	3 11 12 4
TP _{topic-m}	2 8 11 9	2 7 10 11	1 3 12 14	2 9 10 9	- - -	3 7 13 7	3 3 14 10	4 10 10 6	3 3 14 10
TP _{total}	2 6 11 11	3 6 12 9	4 4 12 10	4 11 11 4	3 7 13 7	- - -	0 2 9 19	10 11 7 2	0 2 9 19
TP _{add}	3 4 14 9	2 8 10 10	4 4 11 11	3 11 12 4	3 3 14 10	0 2 9 19	- - -	7 8 11 4	0 0 0 30
TP _{art}	5 11 11 3	6 8 12 4	3 10 11 6	8 11 6 5	4 10 10 6	10 11 7 2	7 8 11 4	- - -	7 8 11 4
TP _{search}	3 3 15 9	2 8 10 10	4 3 13 10	3 11 12 4	3 3 14 10	0 2 9 19	0 0 0 30	7 8 11 4	- - -
Total	22 48 85 85	20 55 83 82	27 37 96 80	35 86 74 45	20 50 94 76	26 49 84 81	22 40 81 97	50 77 79 34	22 38 84 96

TABLE XVI. THE TABLE SHOWS THE CLASSIFICATION OF SUBJECTS AT THE CUT POINT 30% ON TEST-CLASS LEVEL USING JACCARD DISTANCE. THE FOUR VALUES IN EACH CELL ARE THE NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY TWO TECHNIQUES ARE HIGHLY DISSIMILAR, DISSIMILAR, SIMILAR AND HIGHLY SIMILAR RESPECTIVELY. THE LAST ROW SHOWS FOR EACH TECHNIQUE, THE TOTAL NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY THIS TECHNIQUE AND OTHER TECHNIQUES ARE HIGHLY ORTHOGONAL, ORTHOGONAL, SIMILAR, AND HIGHLY SIMILAR RESPECTIVELY.

Techniques	TP _{cg}	TP _{cg-add}	TP _{str}	TP _{topic-r}	TP _{topic-m}	TP _{total}	TP _{add}	TP _{art}	TP _{search}
TP _{cg}	- - - -	0 3 6 21	1 3 11 15	3 14 6 7	2 5 15 8	2 4 13 11	2 3 14 11	5 9 13 3	3 2 14 11
TP _{cg-add}	0 3 6 21	- - - -	1 3 14 12	2 12 9 7	2 7 10 11	2 5 14 9	2 4 13 11	4 8 13 5	2 4 13 11
TP _{str}	1 3 11 15	1 3 14 12	- - - -	6 9 8 7	1 3 10 16	4 3 13 10	4 2 9 15	2 10 11 7	4 2 9 15
TP _{topic-r}	3 14 6 7	2 12 9 7	6 9 8 7	- - - -	2 8 9 11	3 10 12 5	3 9 10 8	4 11 9 6	3 9 10 8
TP _{topic-m}	2 5 15 8	2 7 10 11	1 3 10 16	2 8 9 11	- - - -	3 5 16 6	3 3 13 11	2 11 11 6	4 2 13 11
TP _{total}	2 4 13 11	2 5 14 9	4 3 13 10	3 10 12 5	3 5 16 6	- - - -	0 1 10 19	8 11 8 3	0 1 10 19
TP _{add}	2 3 14 11	2 4 13 11	4 2 9 15	3 9 10 8	3 3 13 11	0 1 10 19	- - - -	6 7 11 6	0 0 0 30
TP _{art}	5 9 13 3	4 8 13 5	2 10 11 7	4 11 9 6	2 11 11 6	8 11 8 3	6 7 11 6	- - - -	6 7 11 6
TP _{search}	3 2 14 11	2 4 13 11	4 2 9 15	3 9 10 8	4 2 13 11	0 1 10 19	0 0 0 30	6 7 11 6	- - - -
Total	18 43 92 87	15 46 92 87	23 35 85 97	26 82 73 59	19 44 97 80	22 40 96 82	20 29 80 111	37 74 87 42	22 27 80 111

TABLE XVII. THE TABLE SHOWS THE CLASSIFICATION OF SUBJECTS AT THE CUT POINT 35% ON TEST-CLASS LEVEL USING JACCARD DISTANCE. THE FOUR VALUES IN EACH CELL ARE THE NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY TWO TECHNIQUES ARE HIGHLY DISSIMILAR, DISSIMILAR, SIMILAR AND HIGHLY SIMILAR RESPECTIVELY. THE LAST ROW SHOWS FOR EACH TECHNIQUE, THE TOTAL NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY THIS TECHNIQUE AND OTHER TECHNIQUES ARE HIGHLY ORTHOGONAL, ORTHOGONAL, SIMILAR, AND HIGHLY SIMILAR RESPECTIVELY.

Techniques	TP _{cg}	TP _{cg-add}	TP _{str}	TP _{topic-r}	TP _{topic-m}	TP _{total}	TP _{add}	TP _{art}	TP _{search}
TP _{cg}	- - - -	0 1 8 21	1 2 8 19	2 13 7 8	2 5 11 12	2 5 12 11	1 1 16 12	2 7 12 9	2 1 15 12
TP _{cg-add}	0 1 8 21	- - - -	2 0 12 16	2 12 8 8	2 4 14 10	2 5 13 10	1 1 14 14	1 6 13 10	1 2 14 13
TP _{str}	1 2 8 19	2 0 12 16	- - - -	4 12 7 7	1 1 11 17	3 3 11 13	2 1 11 16	1 6 12 11	2 2 10 16
TP _{topic-r}	2 13 7 8	2 12 8 8	4 12 7 7	- - - -	2 7 9 12	2 9 11 8	1 6 14 9	2 11 9 8	1 7 13 9
TP _{topic-m}	2 5 11 12	2 4 14 10	1 1 11 17	2 7 9 12	- - - -	2 6 14 8	2 1 15 12	1 9 12 8	2 1 15 12
TP _{total}	2 5 12 11	2 5 13 10	3 3 11 13	2 9 11 8	2 6 14 8	- - - -	1 1 12 16	4 11 10 5	1 1 12 16
TP _{add}	1 1 16 12	1 1 14 14	2 1 11 16	1 6 14 9	2 1 15 12	1 1 12 16	- - - -	2 7 9 12	0 0 1 29
TP _{art}	2 7 12 9	1 6 13 10	1 6 12 11	2 11 9 8	1 9 12 8	4 11 10 5	2 7 9 12	- - - -	2 7 10 11
TP _{search}	2 1 15 12	2 1 14 13	2 2 10 16	1 7 13 9	2 1 15 12	1 1 12 16	0 0 1 29	2 7 10 11	- - - -
Total	12 35 89 104	11 31 96 102	16 27 82 115	16 77 78 69	14 34 101 91	17 41 95 87	10 18 92 120	15 64 87 74	11 21 90 118

TABLE XVIII. THE TABLE SHOWS THE CLASSIFICATION OF SUBJECTS AT THE CUT POINT 40% ON TEST-CLASS LEVEL USING JACCARD DISTANCE. THE FOUR VALUES IN EACH CELL ARE THE NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY TWO TECHNIQUES ARE HIGHLY DISSIMILAR, DISSIMILAR, SIMILAR AND HIGHLY SIMILAR RESPECTIVELY. THE LAST ROW SHOWS FOR EACH TECHNIQUE, THE TOTAL NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY THIS TECHNIQUE AND OTHER TECHNIQUES ARE HIGHLY ORTHOGONAL, ORTHOGONAL, SIMILAR, AND HIGHLY SIMILAR RESPECTIVELY.

Techniques	TP _{cg}	TP _{cg-add}	TP _{str}	TP _{topic-r}	TP _{topic-m}	TP _{total}	TP _{add}	TP _{art}	TP _{search}
TP _{cg}	- - - -	0 2 4 24	1 2 6 21	2 10 10 8	1 2 11 16	2 4 9 15	0 2 12 16	3 4 15 8	0 3 11 16
TP _{cg-add}	0 2 4 24	- - - -	2 0 9 19	1 10 10 9	0 3 10 17	1 4 13 12	0 1 12 17	2 4 14 10	0 2 11 17
TP _{str}	1 2 6 21	2 0 9 19	- - - -	4 8 10 8	1 2 8 19	3 4 9 14	2 0 12 16	0 5 11 14	2 1 11 16
TP _{topic-r}	2 10 10 8	1 10 10 9	4 8 10 8	- - - -	1 7 10 12	2 9 12 7	1 4 14 11	2 6 14 8	1 4 14 11
TP _{topic-m}	1 2 11 16	0 3 10 17	1 2 8 19	1 7 10 12	- - - -	1 6 11 12	1 1 14 14	2 7 11 10	1 1 14 14
TP _{total}	2 4 9 15	1 4 13 12	3 4 9 14	2 9 12 7	1 6 11 12	- - - -	1 2 9 18	3 12 8 7	1 2 9 18
TP _{add}	0 2 12 16	0 1 12 17	2 0 12 16	1 4 14 11	1 1 14 14	1 2 9 18	- - - -	1 6 11 12	0 0 0 30
TP _{art}	3 4 15 8	2 4 14 10	0 5 11 14	2 6 14 8	2 7 11 10	3 12 8 7	1 6 11 12	- - - -	1 6 12 11
TP _{search}	0 3 11 16	0 2 11 17	2 1 11 16	1 4 14 11	1 1 14 14	1 2 9 18	0 0 0 30	1 6 12 11	- - - -
Total	9 29 78 124	6 26 83 125	15 22 76 127	14 58 94 74	8 29 89 114	14 43 80 103	6 16 84 134	14 50 96 80	6 19 82 133

TABLE XIX. THE TABLE SHOWS THE CLASSIFICATION OF SUBJECTS AT THE CUT POINT 45% ON TEST-CLASS LEVEL USING JACCARD DISTANCE. THE FOUR VALUES IN EACH CELL ARE THE NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY TWO TECHNIQUES ARE HIGHLY DISSIMILAR, DISSIMILAR, SIMILAR AND HIGHLY SIMILAR RESPECTIVELY. THE LAST ROW SHOWS FOR EACH TECHNIQUE, THE TOTAL NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY THIS TECHNIQUE AND OTHER TECHNIQUES ARE HIGHLY ORTHOGONAL, ORTHOGONAL, SIMILAR, AND HIGHLY SIMILAR RESPECTIVELY.

Techniques	TP _{cg}	TP _{cg-add}	TP _{str}	TP _{topic-r}	TP _{topic-m}	TP _{total}	TP _{add}	TP _{art}	TP _{search}
TP _{cg}	- - - -	0 2 5 23	1 2 5 22	1 9 13 7	1 2 11 16	1 4 6 19	0 2 11 17	1 4 17 8	0 2 11 17
TP _{cg-add}	0 2 5 23	- - - -	2 0 8 20	0 9 11 10	0 3 11 16	0 2 12 16	0 1 9 20	1 5 13 11	0 1 11 18
TP _{str}	1 2 5 22	2 0 8 20	- - - -	2 8 10 10	1 2 7 20	2 3 9 16	2 0 12 16	0 6 9 15	2 0 12 16
TP _{topic-r}	1 9 13 7	0 9 11 10	2 8 10 10	- - - -	0 6 11 13	0 6 15 9	1 2 14 13	2 4 14 10	1 2 15 12
TP _{topic-m}	1 2 11 16	0 3 11 16	1 2 7 20	0 6 11 13	- - - -	0 5 12 13	1 1 12 16	2 4 13 11	1 1 13 15
TP _{total}	1 4 6 19	0 2 12 16	2 3 9 16	0 6 15 9	0 5 12 13	- - - -	0 1 11 18	1 9 12 8	0 1 12 17
TP _{add}	0 2 11 17	0 1 9 20	2 0 12 16	1 2 14 13	1 1 12 16	0 1 11 18	- - - -	1 5 9 15	0 0 0 30
TP _{art}	1 4 17 8	1 5 13 11	0 6 9 15	2 4 14 10	2 4 13 11	1 9 12 8	1 5 9 15	- - - -	1 5 9 15
TP _{search}	0 2 11 17	0 1 11 18	2 0 12 16	1 2 15 12	1 1 13 15	0 1 12 17	0 0 0 30	1 5 9 15	- - - -
Total	5 27 79 129	3 23 80 134	12 21 72 135	7 46 103 84	6 24 90 120	4 31 89 116	5 12 78 145	9 42 96 93	5 12 83 140

TABLE XX. THE TABLE SHOWS THE CLASSIFICATION OF SUBJECTS AT THE CUT POINT 50% ON TEST-CLASS LEVEL USING JACCARD DISTANCE. THE FOUR VALUES IN EACH CELL ARE THE NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY TWO TECHNIQUES ARE HIGHLY DISSIMILAR, DISSIMILAR, SIMILAR AND HIGHLY SIMILAR RESPECTIVELY. THE LAST ROW SHOWS FOR EACH TECHNIQUE, THE TOTAL NUMBERS OF SUBJECT PROJECTS, THE FAULTS OF WHICH DETECTED BY THIS TECHNIQUE AND OTHER TECHNIQUES ARE HIGHLY ORTHOGONAL, ORTHOGONAL, SIMILAR, AND HIGHLY SIMILAR RESPECTIVELY.

Techniques	TP _{cg}	TP _{cg-add}	TP _{str}	TP _{topic-r}	TP _{topic-m}	TP _{total}	TP _{add}	TP _{art}	TP _{search}
TP _{cg}	- - - -	0 1 5 24	0 1 4 25	1 3 16 10	1 1 10 18	0 3 7 20	0 1 10 19	1 3 14 12	0 1 10 19
TP _{cg-add}	0 1 5 24	- - - -	1 0 6 23	0 5 13 12	0 1 9 20	0 1 11 18	0 1 6 23	1 4 12 13	0 1 6 23
TP _{str}	0 1 4 25	1 0 6 23	- - - -	1 5 9 15	0 2 4 24	1 2 8 19	1 0 6 23	1 4 10 15	1 0 6 23
TP _{topic-r}	1 3 16 10	0 5 13 12	1 5 9 15	- - - -	0 5 10 15	0 4 15 11	0 2 12 16	2 3 10 15	0 2 12 16
TP _{topic-m}	1 1 10 18	0 1 9 20	0 2 4 24	0 5 10 15	- - - -	0 5 7 18	0 2 6 22	2 2 9 17	0 2 7 21
TP _{total}	0 3 7 20	0 1 11 18	1 2 8 19	0 4 15 11	0 5 7 18	- - - -	0 0 10 20	1 7 14 8	0 0 10 20
TP _{add}	0 1 10 19	0 1 6 23	1 0 6 23	0 2 12 16	0 2 6 22	0 0 10 20	- - - -	1 4 7 18	0 0 0 30
TP _{art}	1 3 14 12	1 4 12 13	1 4 10 15	2 3 10 15	2 2 9 17	1 7 14 8	1 4 7 18	- - - -	1 4 6 19
TP _{search}	0 1 10 19	0 1 6 23	1 0 6 23	0 2 12 16	0 2 7 21	0 0 10 20	0 0 0 30	1 4 6 19	- - - -
Total	3 14 76 147	2 14 68 156	6 14 53 167	4 29 97 110	3 20 62 155	2 22 82 134	2 10 57 171	10 31 82 117	2 10 57 171