

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}	- - - -	6 14 24 14	16 20 17 5	19 24 12 3	19 21 16 2	9 14 12 23	11 19 19 9	28 22 5 3	10 19 17 12
TP _{cg-add}	6 14 24 14	- - - -	11 22 22 3	12 28 15 3	9 28 21 0	12 15 21 10	6 20 23 9	24 23 9 2	6 19 21 12
TP _{str}	16 20 17 5	11 22 22 3	- - - -	10 18 22 8	2 8 12 36	12 22 16 8	10 17 26 5	25 24 8 1	10 19 23 6
TP _{topic-r}	19 24 12 3	12 28 15 3	10 18 22 8	- - - -	7 23 23 5	14 23 19 2	12 32 13 1	19 22 14 3	11 29 16 2
TP _{topic-m}	19 21 16 2	9 28 21 0	2 8 12 36	7 23 23 5	- - - -	9 23 18 8	6 21 25 6	21 27 10 0	7 24 23 4
TP _{total}	9 14 12 23	12 15 21 10	12 22 16 8	14 23 19 2	9 23 18 8	- - - -	3 14 23 18	30 19 8 1	4 16 24 14
TP _{add}	11 19 19 9	6 20 23 9	10 17 26 5	12 32 13 1	6 21 25 6	3 14 23 18	- - - -	23 21 13 1	1 5 19 33
TP _{art}	28 22 5 3	24 23 9 2	25 24 8 1	19 22 14 3	21 27 10 0	30 19 8 1	23 21 13 1	- - - -	26 22 9 1
TP _{search}	10 19 17 12	6 19 21 12	10 19 23 6	11 29 16 2	7 24 23 4	4 16 24 14	1 5 19 33	26 22 9 1	- - - -
Total	118 153 122 71	86 169 156 53	96 150 146 72	104 199 134 27	80 175 148 61	93 146 141 84	72 149 161 82	196 180 76 12	75 153 152 84

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}	- - - -	3 14 28 13	11 23 17 7	11 29 15 3	12 20 20 6	6 14 19 19	6 19 22 11	21 23 11 3	3 21 20 14
TP _{cg-add}	3 14 28 13	- - - -	7 18 24 9	5 22 26 5	6 19 28 5	1 21 25 11	2 14 23 19	14 21 19 4	3 14 25 16
TP _{str}	11 23 17 7	7 18 24 9	- - - -	4 21 27 6	0 3 17 38	7 16 23 12	4 12 26 16	15 22 20 1	5 12 29 12
TP _{topic-r}	11 29 15 3	5 22 26 5	4 21 27 6	- - - -	6 22 26 4	7 27 21 3	5 26 23 4	12 25 19 2	4 26 22 6
TP _{topic-m}	12 20 20 6	6 19 28 5	0 3 17 38	6 22 26 4	- - - -	7 17 24 10	6 5 34 13	13 24 20 1	7 7 34 10
TP _{total}	6 14 19 19	1 21 25 11	7 16 23 12	7 27 21 3	7 17 24 10	- - - -	1 11 29 17	19 21 16 2	2 11 26 19
TP _{add}	6 19 22 11	2 14 23 19	4 12 26 16	5 26 23 4	6 5 34 13	1 11 29 17	- - - -	10 19 26 3	1 3 6 48
TP _{art}	21 23 11 3	14 21 19 4	15 22 20 1	12 25 19 2	13 24 20 1	19 21 16 2	10 19 26 3	- - - -	13 17 23 5
TP _{search}	3 21 20 14	3 14 25 16	5 12 29 12	4 26 22 6	7 7 34 10	2 11 26 19	1 3 6 48	13 17 23 5	- - - -
Total	73 163 152 76	41 143 198 82	53 127 183 101	54 198 179 33	57 117 203 87	50 138 183 93	35 109 189 131	117 172 154 21	38 111 185 130

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 12 29 15	9 21 19 9	7 22 25 4	8 18 24 8	5 13 20 20	3 19 21 15	16 22 16 4	2 20 20 16
TP _{cg-add}	2 12 29 15	- - - -	4 18 29 7	3 20 30 5	4 14 28 12	2 19 22 15	2 11 23 22	8 21 22 7	3 14 20 21
TP _{str}	9 21 19 9	4 18 29 7	- - - -	4 16 26 12	0 4 14 40	3 19 24 12	4 8 24 22	6 25 24 3	4 8 27 19
TP _{topic-r}	7 22 25 4	3 20 30 5	4 16 26 12	- - - -	3 16 28 11	4 17 31 6	2 14 36 6	6 22 25 5	2 15 38 3
TP _{topic-m}	8 18 24 8	4 14 28 12	0 4 14 40	3 16 28 11	- - - -	4 15 25 14	5 3 30 20	6 23 26 3	5 4 28 21
TP _{total}	5 13 20 20	2 19 22 15	3 19 24 12	4 17 31 6	4 15 25 14	- - - -	1 9 23 25	16 16 23 3	2 8 22 26
TP _{add}	3 19 21 15	2 11 23 22	4 8 24 22	2 14 36 6	5 3 30 20	1 9 23 25	- - - -	6 20 21 11	0 2 6 50
TP _{art}	16 22 16 4	8 21 22 7	6 25 24 3	6 22 25 5	6 23 26 3	16 16 23 3	6 20 21 11	- - - -	7 20 23 8
TP _{search}	2 20 20 16	3 14 20 21	4 8 27 19	2 15 38 3	5 4 28 21	2 8 22 26	0 2 6 50	7 20 23 8	- - - -
Total	52 147 174 91	28 129 203 104	34 119 187 124	31 142 239 52	35 97 203 129	37 116 190 121	23 86 184 171	71 169 180 44	25 91 184 164

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 6 34 16	6 15 24 13	2 21 28 7	4 13 32 9	2 7 28 21	1 16 25 16	6 22 23 7	1 17 24 16
TP _{cg-add}	2 6 34 16	- - - -	1 13 28 16	2 12 32 12	2 11 26 19	3 14 21 20	1 8 23 26	4 15 25 14	1 8 21 28
TP _{str}	6 15 24 13	1 13 28 16	- - - -	2 14 28 14	2 1 10 45	2 12 26 18	3 5 20 30	4 18 27 9	2 5 22 29
TP _{topic-r}	2 21 28 7	2 12 32 12	2 14 28 14	- - - -	2 12 26 18	3 19 24 12	0 13 35 10	2 17 33 6	0 13 37 8
TP _{topic-m}	4 13 32 9	2 11 26 19	2 1 10 45	2 12 26 18	- - - -	2 11 29 16	3 2 23 30	3 17 30 8	3 1 23 31
TP _{total}	2 7 28 21	3 14 21 20	2 12 26 18	3 19 24 12	2 11 29 16	- - - -	1 8 18 31	5 18 29 6	0 8 19 31
TP _{add}	1 16 25 16	1 8 23 26	3 5 20 30	0 13 35 10	3 2 23 30	1 8 18 31	- - - -	1 14 23 20	0 0 4 54
TP _{art}	6 22 23 7	4 15 25 14	4 18 27 9	2 17 33 6	3 17 30 8	5 18 29 6	1 14 23 20	- - - -	1 13 26 18
TP _{search}	1 17 24 16	1 8 21 28	2 5 22 29	0 13 37 8	3 1 23 31	0 8 19 31	0 0 4 54	1 13 26 18	- - - -
Total	24 117 218 105	16 87 210 151	22 83 185 174	13 121 243 87	21 68 199 176	18 97 194 155	10 66 171 217	26 134 216 88	8 65 176 215

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 7 32 17	2 14 24 18	3 11 38 6	2 10 35 11	1 5 28 24	1 7 33 17	3 17 29 9	1 11 24 22
TP _{cg-add}	2 7 32 17	- - - -	0 11 26 21	1 12 29 16	1 7 28 22	1 13 21 23	0 6 22 30	2 11 31 14	0 7 21 30
TP _{str}	2 14 24 18	0 11 26 21	- - - -	1 11 31 15	0 0 8 50	1 7 24 26	0 4 17 37	3 9 34 12	0 4 19 35
TP _{topic-r}	3 11 38 6	1 12 29 16	1 11 31 15	- - - -	1 9 27 21	1 20 25 12	0 8 38 12	1 10 37 10	0 6 39 13
TP _{topic-m}	2 10 35 11	1 7 28 22	0 0 8 50	1 9 27 21	- - - -	1 7 30 20	1 4 18 35	3 13 29 13	1 3 22 32
TP _{total}	1 5 28 24	1 13 21 23	1 7 24 26	1 20 25 12	1 7 30 20	- - - -	0 4 20 34	3 14 30 11	0 6 14 38
TP _{add}	1 7 33 17	0 6 22 30	0 4 17 37	0 8 38 12	1 4 18 35	0 4 20 34	- - - -	1 5 26 26	0 0 4 54
TP _{art}	3 17 29 9	2 11 31 14	3 9 34 12	1 10 37 10	3 13 29 13	3 14 30 11	1 5 26 26	- - - -	1 6 27 24
TP _{search}	1 11 24 22	0 7 21 30	0 4 19 35	0 6 39 13	1 3 22 32	0 6 14 38	0 0 4 54	1 6 27 24	- - - -
Total	15 82 243 124	7 74 210 173	7 60 183 214	8 87 264 105	10 53 197 204	8 76 192 188	3 38 178 245	17 85 243 119	3 43 170 248

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 8 33 17	2 8 31 17	2 12 35 9	2 5 36 15	0 6 27 25	0 6 32 20	2 14 30 12	0 7 29 22
TP _{cg-add}	0 8 33 17	- - - -	0 5 28 25	0 9 29 20	0 4 30 24	1 9 27 21	0 5 17 36	1 8 29 20	0 5 20 33
TP _{str}	2 8 31 17	0 5 28 25	- - - -	1 7 29 21	0 0 4 54	1 6 22 29	0 3 15 40	3 7 28 20	0 2 18 38
TP _{topic-r}	2 12 35 9	0 9 29 20	1 7 29 21	- - - -	1 7 27 23	0 15 28 15	0 6 33 19	1 7 34 16	0 5 34 19
TP _{topic-m}	2 5 36 15	0 4 30 24	0 0 4 54	1 7 27 23	- - - -	1 5 28 24	0 3 12 43	1 8 29 20	0 3 14 41
TP _{total}	0 6 27 25	1 9 27 21	1 6 22 29	0 15 28 15	1 5 28 24	- - - -	0 2 19 37	2 9 35 12	0 3 16 39
TP _{add}	0 6 32 20	0 5 17 36	0 3 15 40	0 6 33 19	0 3 12 43	0 2 19 37	- - - -	1 2 25 30	0 0 2 56
TP _{art}	2 14 30 12	1 8 29 20	3 7 28 20	1 7 34 16	1 8 29 20	2 9 35 12	1 2 25 30	- - - -	1 2 25 30
TP _{search}	0 7 29 22	0 5 20 33	0 2 18 38	0 5 34 19	0 3 14 41	0 3 16 39	0 0 2 56	1 2 25 30	- - - -
Total	8 66 253 137	2 53 213 196	7 38 175 244	5 68 249 142	5 35 180 244	5 55 202 202	1 27 155 281	12 57 235 160	1 27 158 278

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 33 20	2 6 33 17	1 10 34 13	2 5 34 17	0 5 22 31	0 6 29 23	2 9 33 14	0 6 28 24
TP _{cg-add}	0 5 33 20	- - - -	0 5 23 30	0 5 26 27	0 4 26 28	1 6 24 27	0 4 15 39	0 8 26 24	0 4 17 37
TP _{str}	2 6 33 17	0 5 23 30	- - - -	0 4 28 26	0 0 4 54	1 4 20 33	0 2 13 43	2 5 25 26	0 2 15 41
TP _{topic-r}	1 10 34 13	0 5 26 27	0 4 28 26	- - - -	0 5 24 29	0 8 26 24	0 1 28 29	0 5 25 28	0 1 29 28
TP _{topic-m}	2 5 34 17	0 4 26 28	0 0 4 54	0 5 24 29	- - - -	1 2 25 30	0 3 11 44	1 8 20 29	0 3 13 42
TP _{total}	0 5 22 31	1 6 24 27	1 4 20 33	0 8 26 24	1 2 25 30	- - - -	0 2 14 42	1 9 29 19	0 2 16 40
TP _{add}	0 6 29 23	0 4 15 39	0 2 13 43	0 1 28 29	0 3 11 44	0 2 14 42	- - - -	0 2 18 38	0 0 2 56
TP _{art}	2 9 33 14	0 8 26 24	2 5 25 26	0 5 25 28	1 8 20 29	1 9 29 19	0 2 18 38	- - - -	0 2 18 38
TP _{search}	0 6 28 24	0 4 17 37	0 2 15 41	0 1 29 28	0 3 13 42	0 2 16 40	0 0 2 56	0 2 18 38	- - - -
Total	7 52 246 159	1 41 190 232	5 28 161 270	1 39 220 204	4 30 157 273	4 38 176 246	0 20 130 314	6 48 194 216	0 20 138 306

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 4 29 25	0 5 32 21	1 4 34 19	1 3 29 25	0 2 20 36	0 2 30 26	1 6 35 16	0 2 29 27
TP _{cg-add}	0 4 29 25	- - - -	0 2 19 37	0 2 22 34	0 2 22 34	1 2 24 31	0 2 11 45	0 4 20 34	0 2 11 45
TP _{str}	0 5 32 21	0 2 19 37	- - - -	0 4 27 27	0 1 1 56	0 4 15 39	0 0 9 49	0 4 20 34	0 0 10 48
TP _{topic-r}	1 4 34 19	0 2 22 34	0 4 27 27	- - - -	0 3 21 34	0 6 23 29	0 0 20 38	0 2 21 35	0 0 19 39
TP _{topic-m}	1 3 29 25	0 2 22 34	0 1 1 56	0 3 21 34	- - - -	0 2 19 37	0 1 11 46	0 5 18 35	0 1 11 46
TP _{total}	0 2 20 36	1 2 24 31	0 4 15 39	0 6 23 29	0 2 19 37	- - - -	0 0 14 44	0 5 24 29	0 0 15 43
TP _{add}	0 2 30 26	0 2 11 45	0 0 9 49	0 0 20 38	0 1 11 46	0 0 14 44	- - - -	0 1 11 46	0 0 2 56
TP _{art}	1 6 35 16	0 4 20 34	0 4 20 34	0 2 21 35	0 5 18 35	0 5 24 29	0 1 11 46	- - - -	0 1 10 47
TP _{search}	0 2 29 27	0 2 11 45	0 0 10 48	0 0 19 39	0 1 11 46	0 0 15 43	0 0 2 56	0 1 10 47	- - - -
Total	3 28 238 195	1 20 158 285	0 20 133 311	1 21 187 255	1 18 132 313	1 21 154 288	0 6 108 350	1 28 159 276	0 6 107 351

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 1 25 31	0 3 27 28	1 3 33 21	1 2 29 26	0 1 17 40	0 1 24 33	1 3 32 22	0 2 25 31
TP _{cg-add}	1 1 25 31	- - - -	0 2 14 42	0 2 22 34	0 2 19 37	1 1 21 35	0 1 12 45	0 1 19 38	0 1 12 45
TP _{str}	0 3 27 28	0 2 14 42	- - - -	0 2 23 33	0 1 1 56	0 3 15 40	0 0 11 47	0 2 16 40	0 0 10 48
TP _{topic-r}	1 3 33 21	0 2 22 34	0 2 23 33	- - - -	0 2 20 36	0 2 23 33	0 0 20 38	0 1 17 40	0 0 19 39
TP _{topic-m}	1 2 29 26	0 2 19 37	0 1 1 56	0 2 20 36	- - - -	0 1 20 37	0 0 7 51	0 2 13 43	0 0 8 50
TP _{total}	0 1 17 40	1 1 21 35	0 3 15 40	0 2 23 33	0 1 20 37	- - - -	0 0 12 46	0 1 25 32	0 0 11 47
TP _{add}	0 1 24 33	0 1 12 45	0 0 11 47	0 0 20 38	0 0 7 51	0 0 12 46	- - - -	0 0 6 52	0 0 2 56
TP _{art}	1 3 32 22	0 1 19 38	0 2 16 40	0 1 17 40	0 2 13 43	0 1 25 32	0 0 6 52	- - - -	0 0 7 51
TP _{search}	0 2 25 31	0 1 12 45	0 0 10 48	0 0 19 39	0 0 8 50	0 0 11 47	0 0 2 56	0 0 7 51	- - - -
Total	4 16 212 232	2 11 144 307	0 13 117 334	1 12 177 274	1 10 117 336	1 9 144 310	0 2 94 368	1 10 135 318	0 3 94 367

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 20 37	0 3 20 35	1 1 31 25	1 3 17 37	0 1 13 44	0 1 17 40	1 1 23 33	0 1 17 40
TP _{cg-add}	1 0 20 37	- - - -	0 1 10 47	0 1 16 41	0 0 15 43	0 2 16 40	0 0 9 49	0 0 12 46	0 0 10 48
TP _{str}	0 3 20 35	0 1 10 47	- - - -	0 2 18 38	0 0 3 55	0 1 12 45	0 0 6 52	0 1 10 47	0 0 8 50
TP _{topic-r}	1 1 31 25	0 1 16 41	0 2 18 38	- - - -	0 2 12 44	0 2 20 36	0 0 13 45	0 0 14 44	0 0 14 44
TP _{topic-m}	1 3 17 37	0 0 15 43	0 0 3 55	0 2 12 44	- - - -	0 0 16 42	0 0 5 53	0 1 7 50	0 0 6 52
TP _{total}	0 1 13 44	0 2 16 40	0 1 12 45	0 2 20 36	0 0 16 42	- - - -	0 0 10 48	0 0 19 39	0 0 9 49
TP _{add}	0 1 17 40	0 0 9 49	0 0 6 52	0 0 13 45	0 0 5 53	0 0 10 48	- - - -	0 0 3 55	0 0 2 56
TP _{art}	1 1 23 33	0 0 12 46	0 1 10 47	0 0 14 44	0 1 7 50	0 0 19 39	0 0 3 55	- - - -	0 0 2 56
TP _{search}	0 1 17 40	0 0 10 48	0 0 8 50	0 0 14 44	0 0 6 52	0 0 9 49	0 0 2 56	0 0 2 56	- - - -
Total	4 11 158 291	1 4 108 351	0 8 87 369	1 8 138 317	1 6 81 376	0 6 115 343	0 1 65 398	1 3 90 370	0 1 68 395

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 3 6 49	32 10 9 7	35 14 2 7	31 16 3 8	18 20 10 10	20 20 5 13	37 15 2 4	21 20 4 13
TP _{cg-add}	0 3 6 49	- - - -	31 13 8 6	36 13 2 7	31 15 4 8	20 20 7 11	22 18 5 13	36 17 1 4	22 18 5 13
TP _{str}	32 10 9 7	31 13 8 6	- - - -	38 8 6 6	29 9 8 12	34 14 6 4	36 9 7 6	42 6 2 8	34 11 6 7
TP _{topic-r}	35 14 2 7	36 13 2 7	38 8 6 6	- - - -	10 3 4 41	34 15 5 4	34 16 4 4	35 13 5 5	35 16 3 4
TP _{topic-m}	31 16 3 8	31 15 4 8	29 9 8 12	10 3 4 41	- - - -	31 18 4 5	33 16 4 5	33 17 2 6	33 14 5 6
TP _{total}	18 20 10 10	20 20 7 11	34 14 6 4	34 15 5 4	31 18 4 5	- - - -	0 3 9 46	33 11 6 8	1 4 10 43
TP _{add}	20 20 5 13	22 18 5 13	36 9 7 6	34 16 4 4	33 16 4 5	0 3 9 46	- - - -	34 9 8 7	1 1 2 54
TP _{art}	37 15 2 4	36 17 1 4	42 6 2 8	35 13 5 5	33 17 2 6	33 11 6 8	34 9 8 7	- - - -	34 10 7 7
TP _{search}	21 20 4 13	22 18 5 13	34 11 6 7	35 16 3 4	33 14 5 6	1 4 10 43	1 1 2 54	34 10 7 7	- - - -
Total	194 118 41 111	198 117 38 111	276 80 52 56	257 98 31 78	231 108 34 91	171 105 57 131	180 92 44 148	284 98 33 49	181 94 42 147

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}	- - - -	3 2 16 37	11 13 18 16	28 20 3 7	14 18 17 9	11 19 13 15	15 14 16 13	32 18 5 3	15 15 13 15
TP _{cg-add}	3 2 16 37	- - - -	11 14 15 18	27 23 3 5	13 24 14 7	11 20 14 13	14 16 13 15	30 23 4 1	13 17 10 18
TP _{str}	11 13 18 16	11 14 15 18	- - - -	30 13 10 5	12 12 14 20	20 15 12 11	18 18 13 9	33 15 6 4	18 18 12 10
TP _{topic-r}	28 20 3 7	27 23 3 5	30 13 10 5	- - - -	14 15 12 17	26 17 11 4	24 15 15 4	31 13 7 7	24 16 13 5
TP _{topic-m}	14 18 17 9	13 24 14 7	12 12 14 20	14 15 12 17	- - - -	19 24 7 8	21 20 10 7	30 16 7 5	21 20 9 8
TP _{total}	11 19 13 15	11 20 14 13	20 15 12 11	26 17 11 4	19 24 7 8	- - - -	2 13 11 32	28 13 9 8	2 12 14 30
TP _{add}	15 14 16 13	14 16 13 15	18 18 13 9	24 15 15 4	21 20 10 7	2 13 11 32	- - - -	25 16 13 4	0 0 2 56
TP _{art}	32 18 5 3	30 23 4 1	33 15 6 4	31 13 7 7	30 16 7 5	28 13 9 8	25 16 13 4	- - - -	25 15 15 3
TP _{search}	15 15 13 15	13 17 10 18	18 18 12 10	24 16 13 5	21 20 9 8	2 12 14 30	0 0 2 56	25 15 15 3	- - - -
Total	129 119 101 115	122 139 89 114	153 118 100 93	204 132 74 54	144 149 90 81	119 133 91 121	119 112 93 140	234 129 66 35	118 113 88 145

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}	- - - -	2 4 14 38	7 12 18 21	23 19 7 9	11 15 19 13	7 18 16 17	12 15 18 13	25 22 9 2	11 16 16 15
TP _{cg-add}	2 4 14 38	- - - -	6 14 18 20	23 17 10 8	10 19 17 12	8 15 19 16	11 14 18 15	25 21 7 5	11 14 14 19
TP _{str}	7 12 18 21	6 14 18 20	- - - -	24 12 13 9	9 8 17 24	15 16 16 11	12 21 13 12	25 18 10 5	12 21 15 10
TP _{topic-r}	23 19 7 9	23 17 10 8	24 12 13 9	- - - -	14 16 17 11	19 22 12 5	15 20 17 6	29 19 5 5	16 18 18 6
TP _{topic-m}	11 15 19 13	10 19 17 12	9 8 17 24	14 16 17 11	- - - -	17 18 15 8	15 17 14 12	26 21 8 3	16 16 13 13
TP _{total}	7 18 16 17	8 15 19 16	15 16 16 11	19 22 12 5	17 18 15 8	- - - -	2 8 15 33	25 19 10 4	2 8 17 31
TP _{add}	12 15 18 13	11 14 18 15	12 21 13 12	15 20 17 6	15 17 14 12	2 8 15 33	- - - -	24 17 14 3	0 0 3 55
TP _{art}	25 22 9 2	25 21 7 5	25 18 10 5	29 19 5 5	26 21 8 3	25 19 10 4	24 17 14 3	- - - -	24 17 15 2
TP _{search}	11 16 16 15	11 14 14 19	12 21 15 10	16 18 18 6	16 16 13 13	2 8 17 31	0 0 3 55	24 17 15 2	- - - -
Total	98 121 117 128	96 118 117 133	110 122 120 112	163 143 99 59	118 130 120 96	95 124 120 125	91 112 112 149	203 154 78 29	92 110 111 151

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}	- - - -	1 4 12 41	3 8 21 26	13 24 12 9	4 16 22 16	3 14 23 18	4 17 23 14	21 22 12 3	4 17 24 13
TP _{cg-add}	1 4 12 41	- - - -	3 8 23 24	14 18 16 10	4 14 21 19	4 13 23 18	2 16 25 15	22 17 15 4	2 16 26 14
TP _{str}	3 8 21 26	3 8 23 24	- - - -	17 15 16 10	3 8 22 25	10 10 20 18	8 13 20 17	22 15 13 8	9 12 20 17
TP _{topic-r}	13 24 12 9	14 18 16 10	17 15 16 10	- - - -	11 20 14 13	13 22 15 8	11 23 19 5	21 19 9 9	12 22 19 5
TP _{topic-m}	4 16 22 16	4 14 21 19	3 8 22 25	11 20 14 13	- - - -	11 15 16 16	9 12 22 15	17 26 11 4	10 11 22 15
TP _{total}	3 14 23 18	4 13 23 18	10 10 20 18	13 22 15 8	11 15 16 16	- - - -	1 4 15 38	17 23 14 4	1 5 15 37
TP _{add}	4 17 23 14	2 16 25 15	8 13 20 17	11 23 19 5	9 12 22 15	1 4 15 38	- - - -	15 20 17 6	0 0 1 57
TP _{art}	21 22 12 3	22 17 15 4	22 15 13 8	21 19 9 9	17 26 11 4	17 23 14 4	15 20 17 6	- - - -	15 20 18 5
TP _{search}	4 17 24 13	2 16 26 14	9 12 20 17	12 22 19 5	10 11 22 15	1 5 15 37	0 0 1 57	15 20 18 5	- - - -
Total	53 122 149 140	52 106 161 145	75 89 155 145	112 163 120 69	69 122 150 123	60 106 141 157	50 105 142 167	150 162 109 43	53 103 145 163

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 9 46	2 8 19 29	10 24 14 10	2 17 20 19	2 11 27 18	3 12 26 17	13 24 17 4	3 11 28 16
TP _{cg-add}	0 3 9 46	- - - -	3 5 23 27	9 22 17 10	3 13 20 22	3 13 25 17	2 15 23 18	13 20 19 6	2 15 24 17
TP _{str}	2 8 19 29	3 5 23 27	- - - -	12 16 17 13	3 6 21 28	8 11 17 22	7 9 22 20	13 19 17 9	7 8 24 19
TP _{topic-r}	10 24 14 10	9 22 17 10	12 16 17 13	- - - -	5 17 19 17	10 20 21 7	5 21 26 6	16 21 12 9	5 21 26 6
TP _{topic-m}	2 17 20 19	3 13 20 22	3 6 21 28	5 17 19 17	- - - -	7 17 21 13	5 13 24 16	13 22 16 7	5 13 24 16
TP _{total}	2 11 27 18	3 13 25 17	8 11 17 22	10 20 21 7	7 17 21 13	- - - -	0 5 19 34	15 25 14 4	0 6 18 34
TP _{add}	3 12 26 17	2 15 23 18	7 9 22 20	5 21 26 6	5 13 24 16	0 5 19 34	- - - -	12 21 16 9	0 0 0 58
TP _{art}	13 24 17 4	13 20 19 6	13 19 17 9	16 21 12 9	13 22 16 7	15 25 14 4	12 21 16 9	- - - -	12 21 16 9
TP _{search}	3 11 28 16	2 15 24 17	7 8 24 19	5 21 26 6	5 13 24 16	0 6 18 34	0 0 0 58	12 21 16 9	- - - -
Total	35 110 160 159	35 106 160 163	55 82 160 167	72 162 152 78	43 118 165 138	45 108 162 149	34 96 156 178	107 173 127 57	34 95 160 175

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 13 42	1 8 19 30	7 23 17 11	2 11 27 18	2 8 28 20	2 7 28 21	10 23 20 5	3 6 28 21
TP _{cg-add}	0 3 13 42	- - -	1 6 22 29	6 21 19 12	2 11 22 23	2 14 20 22	2 6 28 22	8 21 20 9	2 6 30 20
TP _{str}	1 8 19 30	1 6 22 29	- - -	9 16 21 12	2 4 21 31	5 10 21 22	6 5 23 24	7 22 19 10	6 5 22 25
TP _{topic-r}	7 23 17 11	6 21 19 12	9 16 21 12	- - -	5 15 20 18	8 17 25 8	6 19 21 12	10 21 16 11	6 19 20 13
TP _{topic-m}	2 11 27 18	2 11 22 23	2 4 21 31	5 15 20 18	- - -	3 15 28 12	4 9 26 19	9 20 20 9	5 8 26 19
TP _{total}	2 8 28 20	2 14 20 22	5 10 21 22	8 17 25 8	3 15 28 12	- - -	0 4 21 33	14 22 17 5	0 4 22 32
TP _{add}	2 7 28 21	2 6 28 22	6 5 23 24	6 19 21 12	4 9 26 19	0 4 21 33	- - -	11 17 21 9	0 0 0 58
TP _{art}	10 23 20 5	8 21 20 9	7 22 19 10	10 21 16 11	9 20 20 9	14 22 17 5	11 17 21 9	- - -	11 17 21 9
TP _{search}	3 6 28 21	2 6 30 20	6 5 22 25	6 19 20 13	5 8 26 19	0 4 22 32	0 0 0 58	11 17 21 9	- - -
Total	27 89 180 168	23 88 174 179	37 76 168 183	57 151 159 97	32 93 190 149	34 94 182 154	31 67 168 198	80 163 154 67	33 65 169 197

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 15 42	1 6 16 35	5 19 21 13	2 10 22 24	2 8 25 23	1 4 31 22	7 20 19 12	2 5 29 22
TP _{cg-add}	0 1 15 42	- - -	2 1 20 35	5 19 19 15	2 7 26 23	2 11 21 24	1 2 30 25	4 19 21 14	1 4 30 23
TP _{str}	1 6 16 35	2 1 20 35	- - -	6 16 23 13	2 2 21 33	4 7 21 26	3 3 21 30	5 18 17 14	4 5 18 31
TP _{topic-r}	5 19 21 13	5 19 19 15	6 16 23 13	- - -	4 13 21 20	5 15 26 12	3 14 25 16	8 19 17 14	3 16 23 16
TP _{topic-m}	2 10 22 24	2 7 26 23	2 2 21 33	4 13 21 20	- - -	2 13 26 17	3 5 28 22	8 17 22 11	3 6 27 22
TP _{total}	2 8 25 23	2 11 21 24	4 7 21 26	5 15 26 12	2 13 26 17	- - -	1 2 21 34	9 22 18 9	1 3 22 32
TP _{add}	1 4 31 22	2 2 30 25	4 3 21 30	3 14 25 16	3 5 28 22	1 2 21 34	- - -	7 14 20 17	0 0 2 56
TP _{art}	7 20 19 12	4 19 21 14	5 18 21 14	8 19 17 14	8 17 22 11	9 22 18 9	7 14 20 17	- - -	7 14 20 17
TP _{search}	2 5 29 22	1 4 30 23	4 5 18 31	3 16 23 16	3 6 27 22	3 5 22 32	0 0 2 56	7 14 20 17	- - -
Total	20 73 178 193	17 64 182 201	28 58 161 217	39 131 175 119	26 73 193 172	26 81 180 177	20 44 178 222	55 143 158 108	21 53 171 219

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 13 43	1 4 13 40	3 16 24 15	1 5 19 33	2 7 19 30	0 5 25 28	7 16 22 13	0 6 24 28
TP _{cg-add}	0 2 13 43	- - -	2 1 16 39	2 16 22 18	0 5 20 33	1 7 24 26	0 1 27 30	3 12 26 17	0 2 27 29
TP _{str}	1 4 13 40	2 1 16 39	- - -	4 11 25 18	1 3 15 39	3 6 19 30	2 1 24 31	2 14 21 21	2 2 23 31
TP _{topic-r}	3 16 24 15	2 16 22 18	4 11 25 18	- - -	1 12 20 25	3 16 27 12	3 6 28 21	4 16 25 13	2 7 27 22
TP _{topic-m}	1 5 19 33	0 5 20 33	1 3 15 39	1 12 20 25	- - -	1 11 25 21	2 3 29 24	5 15 22 16	2 3 28 25
TP _{total}	2 7 19 30	1 7 24 26	3 6 19 30	3 16 27 12	1 11 25 21	- - -	1 3 20 34	5 25 17 11	1 3 20 34
TP _{add}	0 5 25 28	0 1 27 30	2 1 24 31	3 6 28 21	2 3 29 24	1 3 20 34	- - -	3 14 22 19	0 0 0 58
TP _{art}	7 16 22 13	3 12 26 17	2 14 21 21	4 16 25 13	5 15 22 16	5 25 17 11	3 14 22 19	- - -	3 13 24 18
TP _{search}	0 6 24 28	0 2 27 29	2 2 23 31	2 7 27 22	2 3 28 25	1 3 20 34	0 0 0 58	3 13 24 18	- - -
Total	14 61 159 230	8 46 175 235	17 42 156 249	22 100 198 144	13 57 178 216	17 78 171 198	11 33 175 245	32 125 179 128	10 36 173 245

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 12 44	1 4 12 41	2 15 26 15	1 5 19 33	1 7 18 32	0 5 23 30	3 17 25 13	0 5 23 30
TP _{cg-add}	0 2 12 44	- - -	2 1 13 42	1 14 21 22	0 5 21 32	0 3 25 30	0 1 22 35	2 13 23 20	0 1 25 32
TP _{str}	1 4 12 41	2 1 13 42	- - -	2 11 23 22	1 3 12 42	2 5 20 31	2 1 25 30	1 15 20 22	1 1 25 30
TP _{topic-r}	2 15 26 15	1 14 21 22	2 11 23 22	- - -	0 10 22 26	1 12 32 13	3 3 30 22	4 16 22 16	3 3 30 22
TP _{topic-m}	1 5 19 33	0 5 21 32	1 3 12 42	0 10 22 26	- - -	0 10 25 23	2 3 28 25	4 13 24 17	2 3 27 26
TP _{total}	1 7 18 32	0 3 25 30	2 5 20 31	1 12 32 13	0 10 25 23	- - -	0 2 21 35	2 22 22 12	0 2 22 34
TP _{add}	0 5 23 30	0 1 22 35	2 1 25 30	3 3 30 22	2 3 28 25	0 2 21 35	- - -	2 14 20 22	0 0 0 58
TP _{art}	3 17 25 13	2 13 23 20	1 15 20 22	4 16 22 16	4 13 24 17	2 22 22 12	2 14 20 22	- - -	2 12 22 22
TP _{search}	0 5 23 30	0 1 25 32	2 1 25 30	3 3 30 22	2 3 27 26	0 2 22 34	0 0 0 58	2 12 22 22	- - -
Total	8 60 158 238	5 40 162 257	13 41 150 260	16 84 206 158	10 52 178 224	6 63 185 210	9 29 169 257	20 122 178 144	9 27 174 254

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 2 10 46	0 1 10 47	1 6 30 21	1 3 17 37	0 6 17 35	0 2 22 34	2 9 29 18	0 2 22 34
TP _{cg-add}	0 2 10 46	- - -	1 0 9 48	0 6 28 24	0 2 17 39	0 2 23 33	0 1 17 40	2 7 24 25	0 1 16 41
TP _{str}	0 1 10 47	1 0 9 48	- - -	1 6 23 28	0 2 11 45	1 4 19 34	1 1 16 40	2 9 21 26	1 1 16 40
TP _{topic-r}	1 6 30 21	0 6 28 24	1 6 23 28	- - -	0 7 20 31	0 7 32 19	1 3 29 25	4 9 19 26	1 3 29 25
TP _{topic-m}	1 3 17 37	0 2 17 39	0 2 11 45	0 7 20 31	- - -	0 7 20 31	0 4 17 37	3 7 25 23	0 4 18 36
TP _{total}	0 6 17 35	0 2 23 33	1 4 19 34	0 7 32 19	0 7 20 31	- - -	0 2 18 38	2 15 28 13	0 2 18 38
TP _{add}	0 2 22 34	0 1 17 40	1 1 16 40	1 3 29 25	0 4 17 37	0 2 18 38	- - -	2 11 13 32	0 0 0 58
TP _{art}	2 9 29 18	2 7 24 25	2 9 21 26	4 9 19 26	3 7 25 23	2 15 28 13	2 11 13 32	- - -	2 10 13 33
TP _{search}	0 2 22 34	0 1 16 41	1 1 16 40	1 3 29 25	0 4 18 36	0 2 18 38	0 0 0 58	2 10 13 33	- - -
Total	4 31 157 272	3 21 144 296	7 24 125 308	8 47 210 199	4 36 145 279	3 45 175 241	4 24 132 304	19 77 172 196	4 23 132 305