Supplementary material for: "On the supervision of peer assessment tasks: an efficient instructor guidance technique"

Jerónimo Hernández-González[1] and Pedro J. Herrera[2]

[1] *Department of Computer Science, Applied Mathematics and Statistics, University of Girona, Campus Montilivi, 17003 Girona, Spain*
[2] *Department of Software Engineering and Computer Systems, National University of Distance Education, Juan del Rosal 16, 28040 Madrid, Spain*

Table 1: Grades from the real dataset, grouped by peer-assessed student. Each row shows the student id, the peer id: peer-assessed grade ($\times 3$), and the instructor's grade.

Student	Peer #1	Peer #2	Peer #3	Teacher
1	5: 8.00	10: 9.67	15: 8.57	6.90
2	9: 9.67	14: 9.80	15: 8.03	8.73
3	6: 8.93	8: 10.0	14: 9.90	9.03
4	6: 10.0	7: 9.73	11: 9.17	9.47
5	4: 10.0	8: 9.80	11: 9.17	9.77
6	1: 10.0	3: 8.77	9: 9.80	8.93
7	3: 9.27	12: 8.93	16: 9.67	8.73
8	2: 10.0	4: 7.73	13: 8.80	7.93
9	3: 9.30	6: 10.0	12: 9.80	9.37
10	5: 9.47	7: 9.03	16: 10.0	8.40
11	1: 9.00	4: 9.17	13: 9.80	8.70
12	2: 9.50	9: 7.63	15: 6.90	7.67
13	7: 10.0	10: 9.57	16: 9.67	8.33
14	2: 9.67	8: 8.43	10: 9.17	7.83
15	11: 5.83	13: 9.93	14: 10.0	7.40
16	1: 10.0	5: 8.87	12: 8.87	8.13

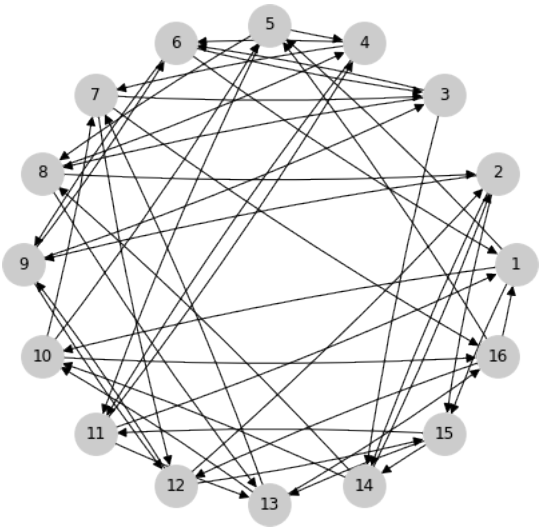

Figure 1: Graph of interconnections among the 16 students. A directed edge $A \rightarrow B$ implies that student A peer-assessed and graded the test of student B.

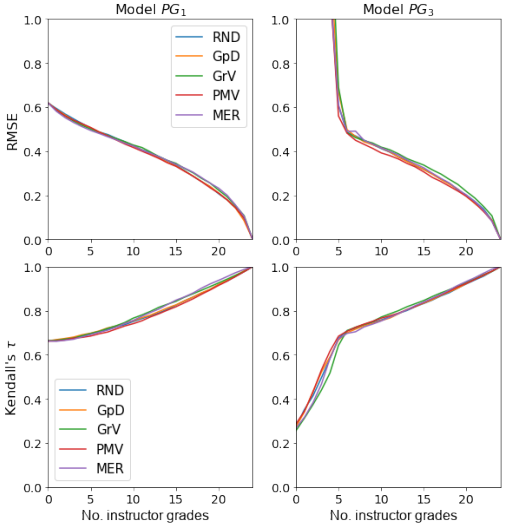

Figure 2: Influence of the selection technique in the performance of the proposal (by rows, in terms of RMSE and *Kendall's* τ) when using the two main models (by columns). Assuming $G = 3$ peer assessments per test, a class of $J = 24$ students and generative model no. 7 (see Table I in the paper), when following different selection criteria (RND, GpD, GrV, PMV or MER), the evolution of the performance is shown as the instructor progresses in the revision of the tests (from none to all of them).