

Rating analysis

December 14, 2023

This notebook analyzes the 3 reviews of the 90 different accepted papers at MICCAI 2023.

1 Categories

For 270 reviews	number	percent	ci low	ci high
Models and algorithms	66	24.44	19.26	29.63
Datasets	91	33.7	28.15	39.26
Code	126	46.67	40.91	52.59
Experimental results	69	25.56	20.74	30.74
Error bars or statistical significance	5	1.85	0.37	3.7
Code is or will be available	62	22.96	18.15	28.15

Table 1: Percent of reviewers that have commented on at least one of the items of the category

2 Statements

For 270 reviews	number	percent	ci low	ci high
3. (+) statement	124	45.93	40	51.5
2. (-) statement	31	11.48	7.41	15.56
1. (none) statement	103	38.15	32.22	43.7
0. Unusable (statement)	12	4.44	2.22	7.04

Table 2: Percent of reviewers that have commented on at least one of the items of the category

kappa fleiss: -0.06679764243614937
confidence intervals (bootstrap): [-0.14%, 0.01%]

2.1 Review 1 VS review 2

	3. (+) statement	2. (-) statement	1. (none) statement	0. Unusable (sta
3. (+) statement	18.89	5.56	18.89	
2. (-) statement	10	0	6.67	
1. (none) statement	16.67	2.22	14.44	
0. Unusable (statement)	0	1.11	2.22	
total	45.56	8.89	42.22	

Table 3: Confusion matrix in percent

2.2 Review 2 VS review 3

	3. (+) statement	2. (-) statement	1. (none) statement	0. Unusable (sta
3. (+) statement	18.89	3.33	20	
2. (-) statement	5.56	1.11	1.11	
1. (none) statement	22.22	4.44	13.33	
0. Unusable (statement)	0	0	3.33	
total	46.67	8.89	37.78	

Table 4: Confusion matrix in percent

2.3 Review 1 VS review 3

	3. (+) statement	2. (-) statement	1. (none) statement	0. Unusable (sta
3. (+) statement	20	6.67	16.67	
2. (-) statement	5.56	1.11	8.89	
1. (none) statement	18.89	1.11	11.11	
0. Unusable (statement)	2.22	0	1.11	
total	46.67	8.89	37.78	

Table 5: Confusion matrix in percent

3 Comments

For 270 reviews	number	percent	ci low	ci high
4. (-/+) comments	62	22.96	18.15	28.17
3. (+) comments	86	31.85	26.09	37.04
2. (-) comments	50	18.52	14.07	23.33
1. (none) comments	60	22.22	17.04	27.04
0. Unusable (comments)	12	4.44	2.22	7.04

Table 6: Percent of reviewers that have commented on at least one of the items of the category

kappa fleiss: 0.04991338241663056
confidence intervals (bootstrap): [-0.03%, 0.13%]

3.1 Review 1 VS review 2

	4. (-/+) comments	3. (+) comments	2. (-) comments	1. (none) comments
4. (-/+) comments	5.56	10	3.33	6.67
3. (+) comments	4.44	12.22	4.44	5.56
2. (-) comments	3.33	6.67	5.56	4.44
1. (none) comments	4.44	5.56	5.56	5.56
0. Unusable (comments)	0	0	2.22	1.11
total	17.78	34.44	21.11	23.33

Table 7: Confusion matrix in percent

3.2 Review 2 VS review 3

	4. (-/+) comments	3. (+) comments	2. (-) comments	1. (none) comments
4. (-/+) comments	7.78	2.22	3.33	3.33
3. (+) comments	6.67	15.56	2.22	7.78
2. (-) comments	4.44	6.67	2.22	5.56
1. (none) comments	4.44	10	3.33	4.44
0. Unusable (comments)	2.22	0	1.11	1.11
total	25.56	34.44	12.22	21.11

Table 8: Confusion matrix in percent

3.3 Review 1 VS review 3

	4. (-/+) comments	3. (+) comments	2. (-) comments	1. (none) comments
4. (-/+) comments	5.56	7.78	3.33	6.67
3. (+) comments	6.67	11.11	3.33	3.33
2. (-) comments	6.67	6.67	3.33	4.44
1. (none) comments	6.67	7.78	2.22	4.44
0. Unusable (comments)	0	1.11	0	2.22
total	25.56	34.44	12.22	21.11

Table 9: Confusion matrix in percent

4 Statements VS comments

in %	4. (-/+) comments	3. (+) comments	2. (-) comments	1. (none) comments
3. (+) statement	19	44	9	
2. (-) statement	10	0	20	
1. (none) statement	33	42	21	
0. Unusable (statement)	0	0	0	
total	62	86	50	

Table 10: Percent of reviewers that have commented on at least one of the items of the category

in %	4. (-/+) comments	3. (+) comments	2. (-) comments	1. (none) comments
3. (+) statement	7.04	16.3	3.33	19
2. (-) statement	3.7	0	7.41	0
1. (none) statement	12.22	15.56	7.78	2
0. Unusable (statement)	0	0	0	
total	22.96	31.85	18.52	22

Table 11: Percent of reviewers that have commented on at least one of the items of the category

5 Meta categories

For 270 reviews	number	percent	ci low	ci high
Reproducible	166	61.48	55.56	66.67
Irreproducible	85	31.48	25.93	37.04
Unusable	19	7.04	4.07	10.37

Table 12: Percent of reviewers that have commented on at least one of the items of the category

kappa fleiss: 0.006038455426664452
confidence intervals (bootstrap): [-0.11%, 0.12%]

5.1 Review 1 VS review 2

	Reproducible	Irreproducible	Unusable	total
Reproducible	36.67	18.89	3.33	58.89
Irreproducible	26.67	8.89	1.11	36.67
Unusable	0	2.22	2.22	4.44
total	63.33	30	6.67	100

Table 13: Confusion matrix in percent

5.2 Review 2 VS review 3

	Reproducible	Irreproducible	Unusable	total
Reproducible	41.11	16.67	5.56	63.33
Irreproducible	18.89	7.78	3.33	30
Unusable	2.22	3.33	1.11	6.67
total	62.22	27.78	10	100

Table 14: Confusion matrix in percent

5.3 Review 1 VS review 3

	Reproducible	Irreproducible	Unusable	total
Reproducible	36.67	17.78	4.44	58.89
Irreproducible	22.22	10	4.44	36.67
Unusable	3.33	0	1.11	4.44
total	62.22	27.78	10	100

Table 15: Confusion matrix in percent

6 Agreement

For 90 reviews	number	percent	ci low	ci high
Agreement	27	30	21.11	38.94
Disagreement	59	65.56	55.56	74.44
Unusable	4	4.44	1.11	8.89

Table 16: Percent of reviewers that have commented on at least one of the items of the category

7 Code

code availableness (in %)	link	N/A	total
0	20	32.22	52.22
1	47.78	0	47.78
total	67.78	32.22	100

Table 17: Percent of reviewers that have commented on at least one of the items of the category

8 Code Review

code availableness (in %)	good link	empty link	no link	total
0	6.67	1.11	5.56	13.33
1	41.11	18.89	26.67	86.67
total	47.78	20	32.22	100

Table 18: Percent of reviewers that have commented on at least one of the items of the category