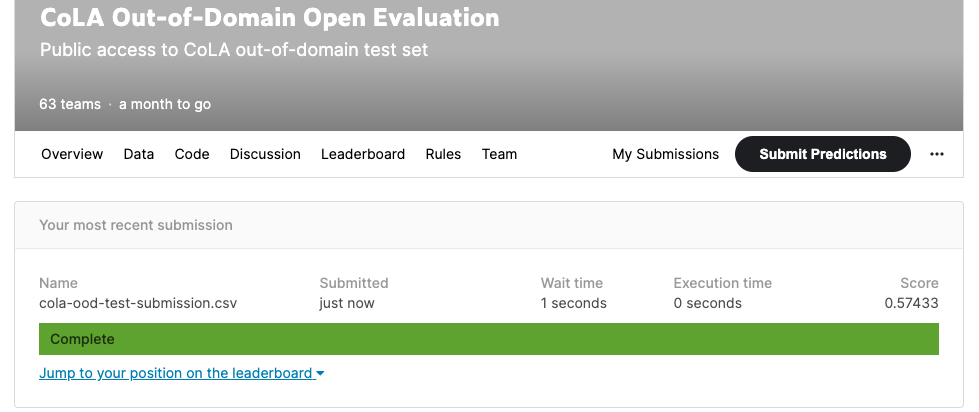
|  | **Rob-cola (matthews corr)** | **Rob-sst2 (acc)** | **Rob-rte (acc)** | **Rob-mnli (acc)** |
| --- | --- | --- | --- | --- |
| **0 shot (base no finetune)** | 0.300 | 0.509 | 0.473 | 0.330 |
| **ID (eval)** | 0.606 | 0.938 | 0.795 | 0.871 |
| **OOD#1** | 0.567 (official cola ood dev, size:516) | 0.876 (10k imdb) | *BUG on mnli* | *BUG on rte* |
| **OOD#2** | 0.574 (kaggle cola ood test, size:533, see Pic 1) | 0.896 (1k yelp) | - | - |
| **OOD#3** | - | 0.906 (38k yelp) | - | - |
| ... |  |  |  |  |
| ... |  |  |  |  |



Pic 1: Kaggle submission result for CoLA-OOD-Test, score is Matthews