

eRisk 2019: T2 results

No Author Given

No Institute Given

1 Task 2: Early Detection of Signs of Self-Harm

This is a new task in 2019. It has the same format as T1 (but T2 provided no training data!).

The challenge consists of sequentially processing pieces of evidence and detect early traces of self-harm as soon as possible. There are two categories of users, self-harm and non-self-harm, and, for each user, the collection contains a sequence of writings (in chronological order). T2 had only a test stage (no training stage) and, therefore, we encouraged participants to design their own unsupervised (e.g. search-based) strategies to detect possible cases of self-harm. The test stage consisted of a period of time where the participants had to connect to our server and iteratively get user writings and send responses.

2 Task 2: evaluation metrics

same as T1's metrics (see T1's evaluation report)

3 Task 2: results

Table 1 shows the participating teams, the number of runs submitted, and the approximate lapse of time from the first response to the last response. Most of the submitted runs processed the entire thread of messages (around 2000 iterations), but a few systems opted for stopping earlier. Compared with T1, the teams were quicker at processing the entire thread of user writings but there were still some teams that took more than a day for running the whole estimation process. Again, this suggests that some participants incorporated some form of offline processing.

team	#runs	#user writings processed	lapse of time (from 1st to last response)
BiTeM	5	8	3 min
BioInfo@UAVR	1	1992	4 hs
Fazl	3	1993	18 days + 21 hs
UNSL	5	1992	13 hs
UDE	5	1992	1 day + 2 hs
LTL-INAOE	4	1993	17 hs
lirmm	5	2004	2 days + 22 hs
CAMH	5	1992	1 day + 19 hs

Table 1. Participating teams: number of runs, number of user writings processed by the team, and lapse of time taken for the whole process.

team	run	<i>P</i>	<i>R</i>	<i>F1</i>	<i>ERDE</i> ₅	<i>ERDE</i> ₅₀	<i>latency</i> _{TP}	<i>speed</i>	<i>latency-weighted F1</i>
BiTeM	0	.52	.41	.46	.10	.08	3	.99	.46
BiTeM	1	1	.05	.09	.12	.11	6.5	.98	.09
BiTeM	2	0	0	0	-	-	-	-	-
BiTeM	3	0	0	0	-	-	-	-	-
BiTeM	4	0	0	0	-	-	-	-	-
BioInfo@UAVR	0	.55	.39	.46	.11	.08	6	.98	.45
Fazl	0	.12	1	.22	.23	.17	51	.81	.17
Fazl	1	.12	1	.22	.23	.16	47	.82	.18
Fazl	2	.12	1	.22	.23	.13	35	.87	.19
UNSL	0	.71	.41	.52	.09	.07	2	1	.52
UNSL	1	.70	.39	.50	.09	.08	2.5	.99	.50
UNSL	2	.20	.90	.32	.09	.07	2	1	.32
UNSL	3	.31	.85	.45	.09	.05	3	.99	.45
UNSL	4	.31	.88	.46	.08	.05	3	.99	.45
UDE	0	.50	.07	.13	.12	.11	13	.95	.12
UDE	1	.45	.22	.30	.11	.10	7	.98	.29
UDE	2	.18	.68	.29	.14	.10	13.5	.95	.28
UDE	3	.06	.34	.10	.20	.20	73.5	.72	.07
UDE	4	0	0	0	-	-	-	-	-
LTL-INAOE	0	.12	1	.22	.13	.11	1	1	.22
LTL-INAOE	1	.12	1	.22	.13	.11	2	1	.21
LTL-INAOE	2	.12	1	.22	.17	.11	4	.99	.21
LTL-INAOE	3	.12	1	.22	.17	.11	5	.98	.21
lirmm	0	.57	.29	.39	.12	.10	28.5	.89	.35
lirmm	1	.53	.22	.31	.12	.11	21	.92	.29
lirmm	2	.48	.49	.48	.13	.12	2004	3e ⁻⁷	1e ⁻⁷
lirmm	3	.47	.44	.46	.13	.12	2004	3e ⁻⁷	1e ⁻⁷
lirmm	4	.52	.41	.46	.13	.12	2004	3e ⁻⁷	1e ⁻⁷
CAMH	0	.12	.95	.22	.17	.11	7	.98	.22
CAMH	1	.12	.93	.22	.17	.11	7	.98	.21
CAMH	2	.12	.90	.22	.17	.11	8	.97	.21
CAMH	3	.12	.98	.22	.16	.11	3.5	.99	.22
CAMH	4	.12	1	.22	.15	.10	4	.99	.22

Table 2. Decision-based evaluation

team	run	1 writing			100 writings			500 writings			1000 writings		
		P@10	NDCG@10	NDCG@100	P@10	NDCG@10	NDCG@100	P@10	NDCG@10	NDCG@100	P@10	NDCG@10	NDCG@100
BiTeM	0	.3	.35	.53	-	-	-	-	-	-	-	-	-
BiTeM	1	.4	.47	.39	-	-	-	-	-	-	-	-	-
BiTeM	2	.5	.48	.44	-	-	-	-	-	-	-	-	-
BiTeM	3	.2	.38	.41	-	-	-	-	-	-	-	-	-
BiTeM	4	.4	.56	.50	-	-	-	-	-	-	-	-	-
BioInfo@UAVR	0	-	-	-	-	-	-	-	-	-	-	-	-
Fazl	0	.1	.12	.30	.1	.06	.42	.1	.06	.41	.6	.40	.59
Fazl	1	.2	.27	.36	.9	.94	.83	.9	.94	.84	.9	.94	.84
Fazl	2	0	0	.13	.6	.73	.73	.7	.68	.71	.7	.68	.74
UNSL	0	.7	.79	.48	.9	.94	.61	.9	.94	.66	.9	.94	.66
UNSL	1	.6	.74	.48	.9	.94	.60	.9	.94	.65	.9	.94	.65
UNSL	2	.9	.88	.62	.8	.75	.75	.5	.59	.74	.6	.64	.74
UNSL	3	1	1	.67	.9	.94	.84	.7	.63	.75	.7	.63	.75
UNSL	4	1	1	.64	.9	.93	.86	.7	.67	.79	.8	.74	.78
UDE	0	0	0	.09	.7	.77	.69	.7	.67	.69	.7	.67	.70
UDE	1	.7	.56	.52	.7	.66	.69	.8	.75	.74	.8	.75	.74
UDE	2	.5	.63	.53	.5	.56	.64	.6	.66	.68	.6	.65	.67
UDE	3	.2	.19	.21	0	0	.11	.2	.16	.14	.1	.07	.15
UDE	4	.2	.25	.30	.1	.07	.20	.1	.07	.15	.1	.08	.17
LTL-INAOE	0	.5	.50	.45	.4	.41	.55	.2	.19	.25	.1	.19	.37
LTL-INAOE	1	.6	.73	.56	.1	.19	.17	.1	.06	.07	0	0	.04
LTL-INAOE	2	.4	.42	.32	.1	.07	.43	.2	.19	.25	.1	.19	.37
LTL-INAOE	3	.7	.72	.44	.1	.19	.27	.1	.06	.19	0	0	.29
lirmm	0	.1	.19	.15	0	0	.01	-	-	-	-	-	-
lirmm	1	.1	.19	.15	0	0	.01	-	-	-	-	-	-
lirmm	2	.1	.19	.15	0	0	.01	-	-	-	-	-	-
lirmm	3	.1	.19	.15	0	0	.01	-	-	-	-	-	-
lirmm	4	.1	.19	.15	0	0	.01	-	-	-	-	-	-
CAMH	0	.3	.37	.47	.6	.71	.49	.7	.72	.50	.6	.66	.48
CAMH	1	.4	.41	.43	.6	.65	.42	.7	.72	.49	.6	.66	.47
CAMH	2	.3	.41	.51	.5	.62	.39	.7	.72	.45	.6	.66	.44
CAMH	3	.3	.25	.42	.5	.55	.32	.7	.72	.37	.6	.66	.37
CAMH	4	.5	.48	.50	.6	.59	.34	.6	.66	.43	.6	.66	.39

Table 3. Ranking-based evaluation