

NLP Results

CMI - 0

BERT

Epoch	Training Loss	Validation Loss	Accuracy
1	No log	1.592136	0.250000
2	1.617200	1.591661	0.250000
3	1.617200	1.590711	0.250000
4	1.600700	1.589285	0.250000
5	1.600700	1.587381	0.250000
6	1.563700	1.584967	0.250000
7	1.563700	1.582105	0.250000
8	1.571300	1.578863	0.250000
9	1.571300	1.574997	0.250000
10	1.483600	1.570830	0.250000

{'eval_loss': 1.5921356678009033, 'eval_accuracy': 0.25, 'eval_runtime': 0.0337, 'eval_samples_per_second': 118.646, 'eval_steps_per_second': 29.661, 'epoch': 10.0}

Muril BERT

Epoch	Training Loss	Validation Loss	Accuracy
1	No log	1.602996	0.750000
2	1.602000	1.602990	0.750000
3	1.602000	1.602978	0.750000
4	1.602900	1.602959	0.750000
5	1.602900	1.602934	0.750000
6	1.602600	1.602903	0.750000
7	1.602600	1.602865	0.750000
8	1.602000	1.602821	0.750000
9	1.602000	1.602771	0.750000
10	1.602100	1.602714	0.750000

{'eval_loss': 1.6029964685440063, 'eval_accuracy': 0.75, 'eval_runtime': 0.0318, 'eval_samples_per_second': 125.841, 'eval_steps_per_second': 31.46, 'epoch': 10.0}

CMI - 0.2

BERT

Epoch	Training Loss	Validation Loss	Accuracy
1	No log	1.579296	0.250000
2	1.619400	1.578784	0.250000
3	1.619400	1.577822	0.250000
4	1.592900	1.576368	0.250000
5	1.592900	1.574174	0.250000
6	1.572400	1.571026	0.250000
7	1.572400	1.567022	0.250000
8	1.588000	1.561674	0.250000
9	1.588000	1.554101	0.250000
10	1.523900	1.542075	0.250000

{'eval_loss': 1.5792959928512573, 'eval_accuracy': 0.25, 'eval_runtime': 0.029, 'eval_samples_per_second': 137.861, 'eval_steps_per_second': 34.465, 'epoch': 10.0}

Muril BERT

Epoch	Training Loss	Validation Loss	Accuracy
1	No log	1.602895	0.750000
2	1.601900	1.602889	0.750000
3	1.601900	1.602878	0.750000
4	1.602900	1.602860	0.750000
5	1.602900	1.602837	0.750000
6	1.602500	1.602808	0.750000
7	1.602500	1.602773	0.750000
8	1.602000	1.602732	0.750000
9	1.602000	1.602686	0.750000
10	1.602100	1.602633	0.750000

{'eval_loss': 1.6028951406478882, 'eval_accuracy': 0.75, 'eval_runtime': 0.0279, 'eval_samples_per_second': 143.567, 'eval_steps_per_second': 35.892, 'epoch': 10.0}

CMI - 0.4

BERT

Epoch	Training Loss	Validation Loss	Accuracy
1	No log	1.581280	0.250000
2	1.583200	1.579646	0.250000
3	1.583200	1.576250	0.250000
4	1.608700	1.571183	0.250000
5	1.608700	1.564315	0.250000
6	1.562500	1.555083	0.250000
7	1.562500	1.544068	0.250000
8	1.547900	1.507101	0.500000
9	1.547900	1.467402	0.500000
10	1.509600	1.452946	0.500000

{'eval_loss': 1.5071009397506714, 'eval_accuracy': 0.5, 'eval_runtime': 0.033,
'eval_samples_per_second': 121.075, 'eval_steps_per_second': 30.269, 'epoch': 10.0}

Muril BERT

Epoch	Training Loss	Validation Loss	Accuracy
1	No log	1.602967	0.750000
2	1.601900	1.602962	0.750000
3	1.601900	1.602950	0.750000
4	1.602900	1.602933	0.750000
5	1.602900	1.602909	0.750000
6	1.602500	1.602880	0.750000
7	1.602500	1.602845	0.750000
8	1.602000	1.602805	0.750000
9	1.602000	1.602758	0.750000
10	1.602100	1.602706	0.750000

{'eval_loss': 1.602967381477356, 'eval_accuracy': 0.75, 'eval_runtime': 0.0285,
'eval_samples_per_second': 140.305, 'eval_steps_per_second': 35.076, 'epoch': 10.0}

CMI - 0.6

BERT

Epoch	Training Loss	Validation Loss	Accuracy
1	No log	1.516126	0.500000
2	1.583500	1.515986	0.500000
3	1.583500	1.514844	0.500000
4	1.580500	1.511901	0.500000
5	1.580500	1.507185	0.500000
6	1.599000	1.500438	0.500000
7	1.599000	1.492116	0.750000
8	1.569100	1.482010	0.750000
9	1.569100	1.457773	0.750000
10	1.542800	1.429417	0.750000

{'eval_loss': 1.4921162128448486, 'eval_accuracy': 0.75, 'eval_runtime': 0.0314, 'eval_samples_per_second': 127.352, 'eval_steps_per_second': 31.838, 'epoch': 10.0}

Muril BERT

Epoch	Training Loss	Validation Loss	Accuracy
1	No log	1.603040	0.750000
2	1.602000	1.603034	0.750000
3	1.602000	1.603022	0.750000
4	1.603000	1.603004	0.750000
5	1.603000	1.602980	0.750000
6	1.602600	1.602950	0.750000
7	1.602600	1.602914	0.750000
8	1.602100	1.602873	0.750000
9	1.602100	1.602825	0.750000
10	1.602200	1.602771	0.750000

{'eval_loss': 1.6030398607254028, 'eval_accuracy': 0.75, 'eval_runtime': 0.0327, 'eval_samples_per_second': 122.456, 'eval_steps_per_second': 30.614, 'epoch': 10.0}

Insights:

From the results it is clear that Muril BERT performs better than BERT in all cases except CMI - 0.6 where the eval loss is lower for BERT for the same accuracy

We can also observe that BERT starts to perform better as the CMI is increased
The accuracy improves from 0.25 to 0.75 from 0 to 0.6 CMI

The accuracy of Muril BERT remains consistent across all CMI