

bert

single-task-score

	Accuracy	Precision	Recall	<u>F1</u>	<u>AUC</u>
aggression	0.924	0.788	0.785	0.787	0.870
attack	0.937	0.825	0.784	0.803	0.875
toxicity	0.937	0.811	0.838	0.824	0.898

multi-task-score

	Accuracy	Precision	Recall	<u>F1</u>	<u>AUC</u>
aggression	0.932	0.823	0.789	0.805	0.876
attack	0.939	0.804	0.830	0.817	0.895
toxicity	0.942	0.845	0.822	0.833	0.895

aggression+attack

	Accuracy	Precision	Recall	<u>F1</u>	<u>AUC</u>
aggression	0.929	0.802	0.800	0.801	0.879
attack	0.935	0.783	0.842	0.811	0.898

aggression+toxicity

	Accuracy	Precision	Recall	<u>F1</u>	<u>AUC</u>
aggression	0.931	0.813	0.797	0.805	0.879
toxicity	0.943	0.840	0.834	0.837	0.900

attack+toxicity

	Accuracy	Precision	Recall	<u>F1</u>	<u>AUC</u>
attack	0.940	0.816	0.820	0.818	0.892
toxicity	0.944	0.861	0.815	0.837	0.893

xlnet

single-task-score

	Accuracy	Precision	Recall	<u>F1</u>	<u>AUC</u>
aggression	0.927	0.802	0.785	0.793	0.872
attack	0.938	0.859	0.750	0.801	0.863
toxicity	0.939	0.829	0.821	0.825	0.893

multi-task-score

	Accuracy	Precision	Recall	<u>F1</u>	<u>AUC</u>
aggression	0.931	0.836	0.764	0.798	0.866
attack	0.939	0.831	0.792	0.811	0.880
toxicity	0.941	0.860	0.794	0.826	0.883

roberta

single-task-score

	Accuracy	Precision	Recall	<u>F1</u>	<u>AUC</u>
aggression	0.914	0.801	0.670	0.730	0.817
attack	0.918	0.787	0.690	0.735	0.827
toxicity	0.927	0.876	0.679	0.765	0.829

multi-task-score

	Accuracy	Precision	Recall	<u>F1</u>	<u>AUC</u>
aggression	0.917	0.812	0.693	0.748	0.829
attack	0.926	0.813	0.713	0.760	0.840
toxicity	0.927	0.830	0.725	0.776	0.847