

[]



Pyth

...

	Metric	Naive	BM25	MultiQ	Parent	Ensemble	Semantic
0	faithfulness	0.886742	0.624167	0.728294	0.846071	0.840016	0.783260
1	answer_relevancy	0.922323	0.777854	0.830801	0.877423	0.816261	0.865046
2	context_recall	0.975000	0.808333	0.925000	0.825000	0.958333	0.875000
3	context_precision	0.761589	0.781944	0.671684	0.787500	0.733769	0.752044
4	answer_correctness	0.665729	0.633674	0.600706	0.626422	0.602225	0.595098