

✓ 0.03							
	Metric	Naive	BM25	MultiQ	Parent	Ensemble	Semantic
0	faithfulness	0.803968	0.625714	0.931944	0.832857	0.781667	0.773268
1	answer_relevancy	0.927901	0.815506	0.875711	0.964192	0.818980	0.919837
2	context_recall	0.975000	0.804167	0.975000	0.875000	0.958333	0.875000
3	context_precision	0.769913	0.684722	0.705357	0.837500	0.672808	0.753054
4	answer_correctness	0.669522	0.566215	0.653025	0.699029	0.664311	0.564746