✓ 0.0s

	Metric	Naive	BM25	MultiQ	Parent	Ensemble	Semantic
0	faithfulness	0.839583	0.673571	0.774325	0.799859	0.874269	0.886250
1	answer_relevancy	0.919344	0.869988	0.818231	0.919560	0.870616	0.935787
2	context_recall	0.975000	0.795833	0.958333	0.875000	0.975000	0.875000
3	context_precision	0.706679	0.709722	0.671226	0.837500	0.693189	0.790545
4	answer_correctness	0.654593	0.613809	0.691301	0.665133	0.716585	0.613560