✓ 0.0s

	Metric	Naive	BM25	MultiQ	Parent	Ensemble	Semantic
0	faithfulness	0.883333	0.663450	0.831389	0.825253	0.825909	0.881429
1	answer_relevancy	0.870336	0.774787	0.780955	0.921902	0.920524	0.873543
2	context_recall	0.975000	0.808333	0.958333	0.875000	0.975000	0.875000
3	context_precision	0.761515	0.736111	0.633504	0.787500	0.676110	0.734071
4	answer_correctness	0.654338	0.628899	0.674448	0.663628	0.705537	0.575849