

Стр. 1 из 2 27.07.2025, 19:47

	sample_id	policy_name	benchmark	first_token_latency	throughput	f1	exact_match	rouge_l
0	auto_00000	early_exit_28	crmarena	0.2856	289.3975	0.3773	0	0.4025
1	auto_00001	early_exit_28	crmarena	0.2687	351.855	0.2663	0	0.3024
2	auto_00002	early_exit_28	crmarena	0.2803	324.4893	0.4817	0	0.504
3	auto_00003	early_exit_28	crmarena	0.2498	268.2883	0.3512	0	0.3776
4	auto_00004	early_exit_28	crmarena	0.1797	329.6628	0.5774	0	0.5958
5	auto_00005	early_exit_28	crmarena	0.1933	322.9427	0.336	0	0.3478
6	auto_00006	early_exit_28	crmarena	0.1965	272.55	0.1968	0	0.2047
7	auto_00007	early_exit_28	crmarena	0.3111	294.5844	0.6087	0	0.6352
8	auto_00008	early_exit_28	crmarena	0.2594	294.9284	0.4371	0	0.4631
9	auto_00009	early_exit_28	crmarena	0.2185	314.6904	0.4033	0	0.3899



Download Filtered Results

Built with 💜 using Streamlit | Last updated: 2025-07-27 19:39:30

Стр. 2 из 2 27.07.2025, 19:47