

	NQ 40k			TriviaQA 40k		
	H@1	H@5	M@5	H@1	H@5	M@5
BGE-FT	53.42	80.15	63.99	52.70	75.22	61.65
BGE-FT + BGE-Reranker	21.91	54.58	33.33	45.36	72.16	55.78
BGE-FT + BGE-Reranker-FT	52.15	76.15	61.37	44.87	67.39	53.39
BGE-FT + RankGPT	44.21	73.68	55.51	48.00	72.00	57.33
GTR-XL	37.64	66.84	48.94	35.97	63.75	46.67
GTR-XL + BGE-Reranker	26.39	59.96	38.50	42.41	68.42	52.51
GTR-XL + BGE-Reranker-FT	57.50	78.92	66.06	58.56	77.65	66.22
GTR-XL + RankGPT	42.11	68.42	52.30	47.00	66.00	54.95
GritLM	44.67	76.00	57.03	39.91	69.34	51.14
GritLM + BGE-Reranker	30.06	65.87	43.20	43.64	70.87	54.23
GritLM + BGE-Reranker-FT	57.57	81.35	66.98	58.60	80.54	67.21
GritLM + RankGPT	37.89	70.53	51.19	44.00	66.00	52.70
DSI-XL	43.03	60.26	49.47	29.64	46.74	36.12
DSI-XL + BGE-Reranker	34.39	64.26	45.74	37.85	52.57	43.49
DSI-XL + BGE-Reranker-FT	50.02	68.60	57.43	36.49	52.40	42.36
DSI-XL + RankGPT	49.47	73.68	59.25	39.00	52.00	44.75
Self-Retrieval (StableLM)	62.16	79.28	69.45	58.69	78.39	66.72
Self-Retrieval (Llama 2)	63.44	79.29	70.00	59.94	81.06	68.74

Table 1: The results of incorporating various rerankers into strong retrieval baselines. We used three different reranker baselines: BGE-reranker, RankGPT, and BGE-reranker-FT.

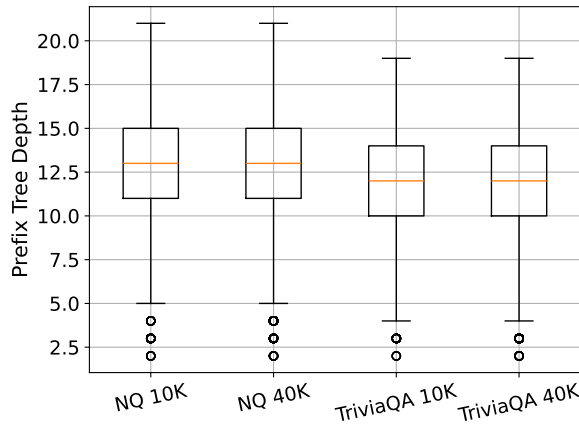


Figure 1: The depth of prefix tree used in our method.