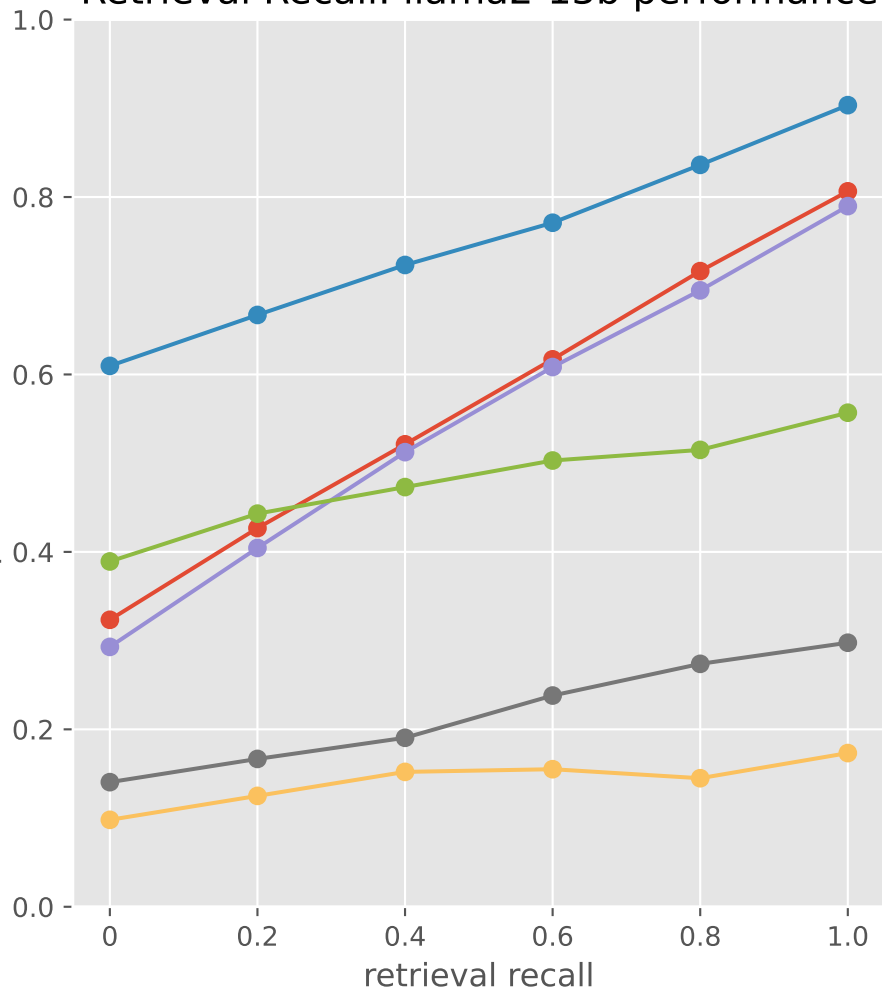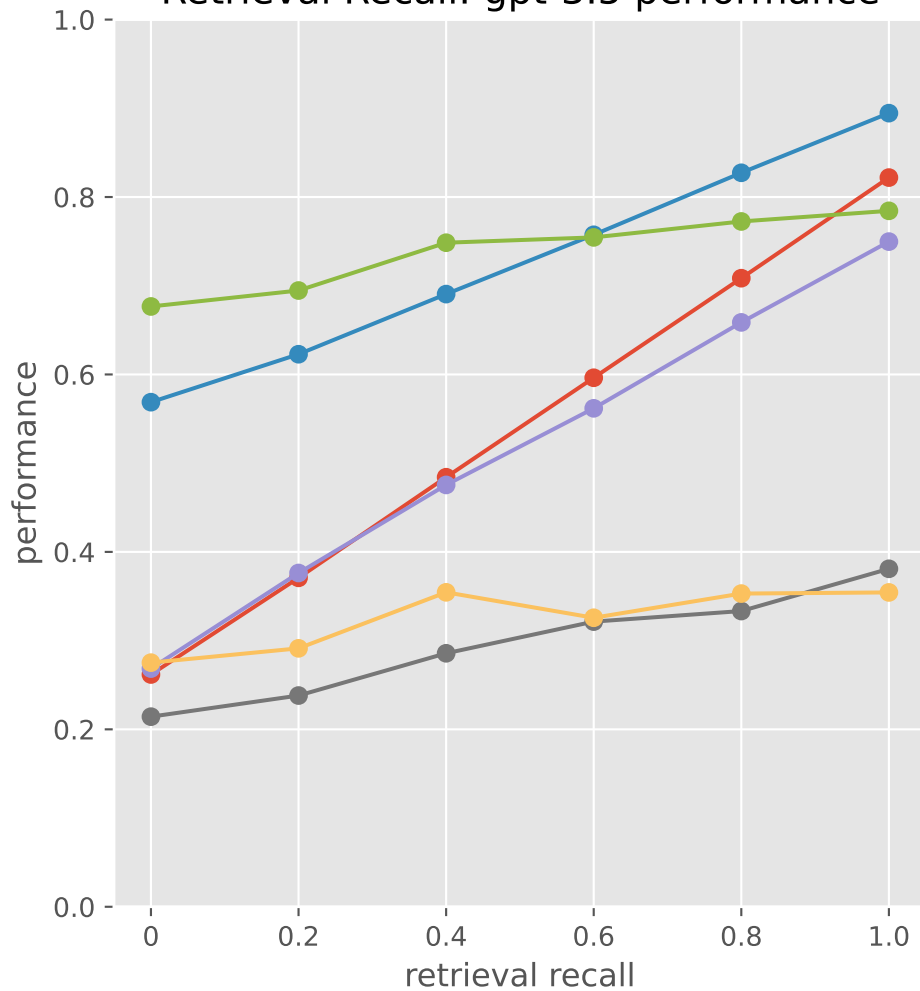Retrieval Recall: llama2-13b performance | Retrieval Recall: gpt-3.5 performance

Legend: NQ, TriviaQA, hotpotQA, conala, DS1000, pandas_numpy_eval