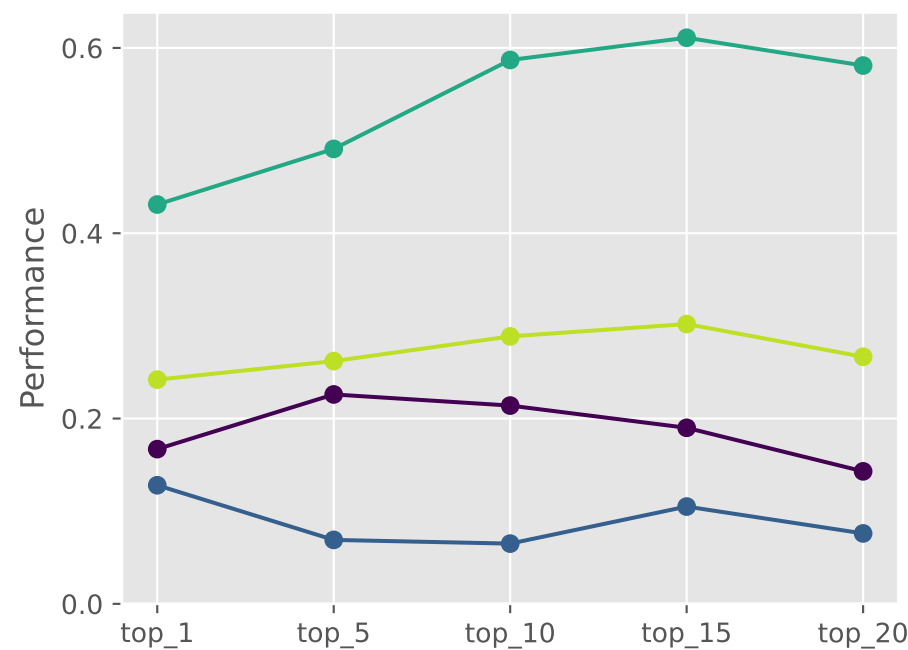
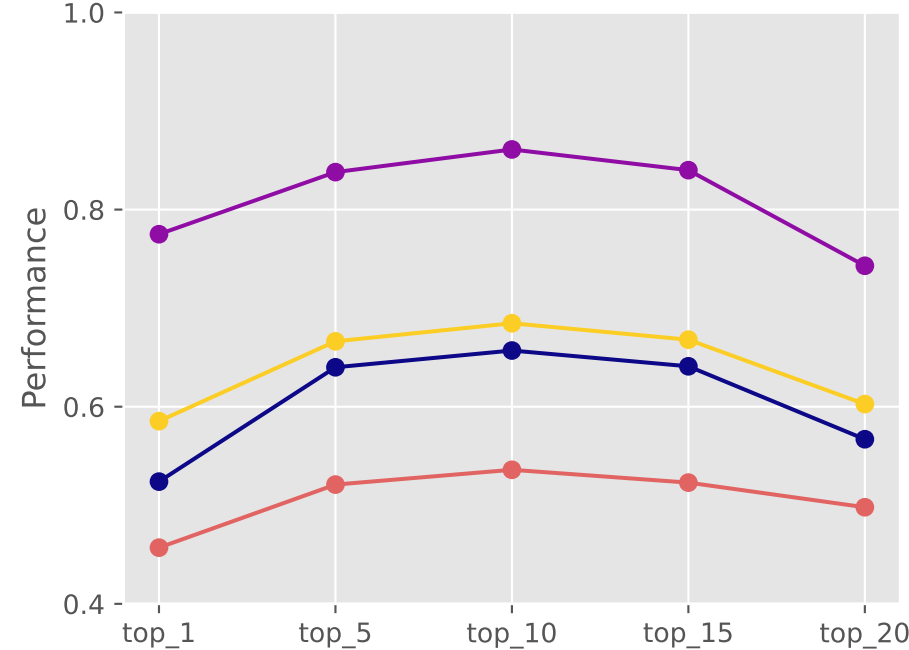


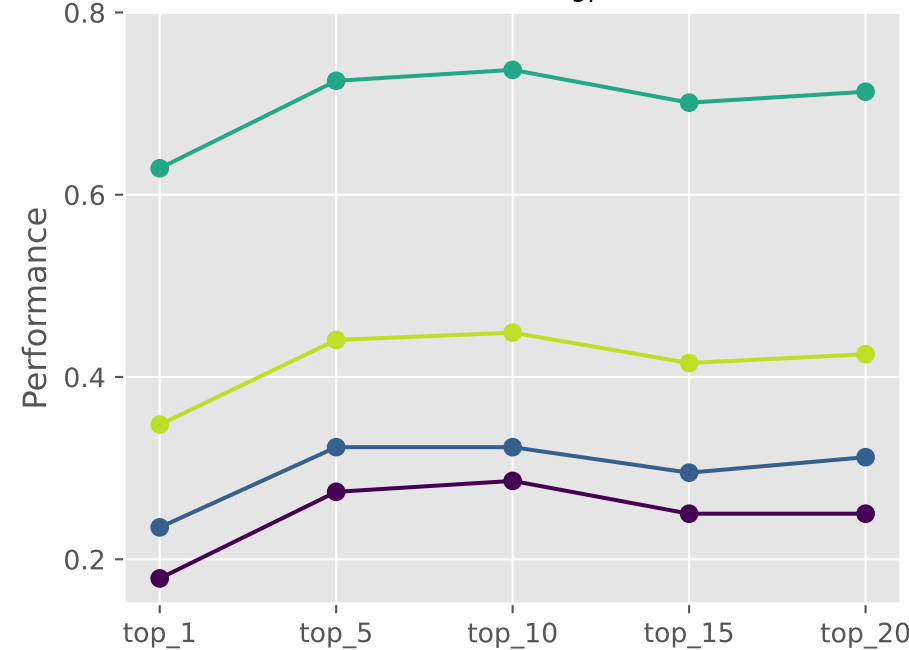
Code Datasets, llama2-13b



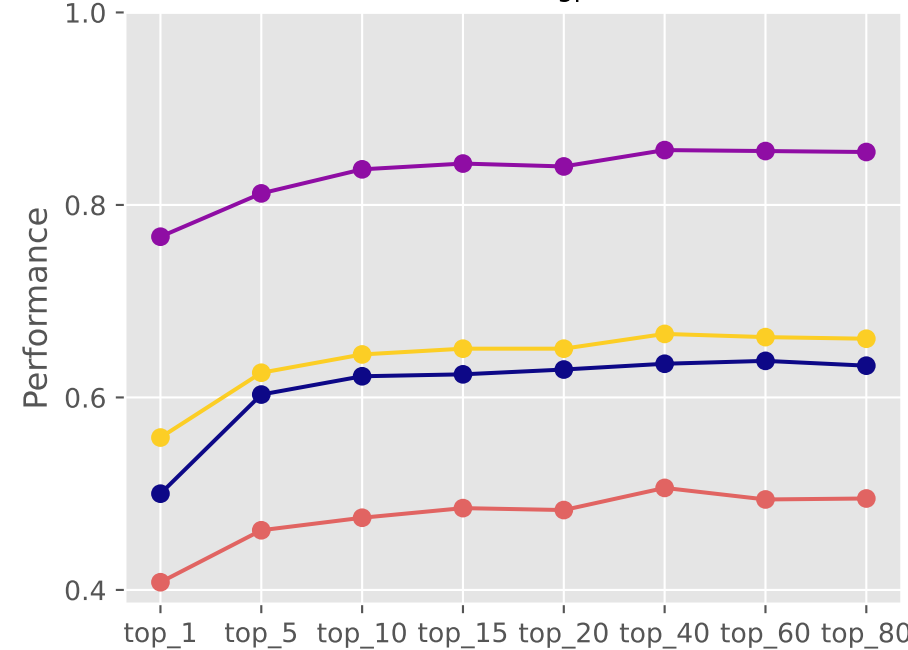
QA Datasets, llama2-13b



Code Datasets, gpt-3.5



QA Datasets, gpt-3.5



Legend:

- conala
- DS1000
- pandas_numpy_eval
- code avg
- NQ
- TriviaQA
- hotpotQA
- qa avg