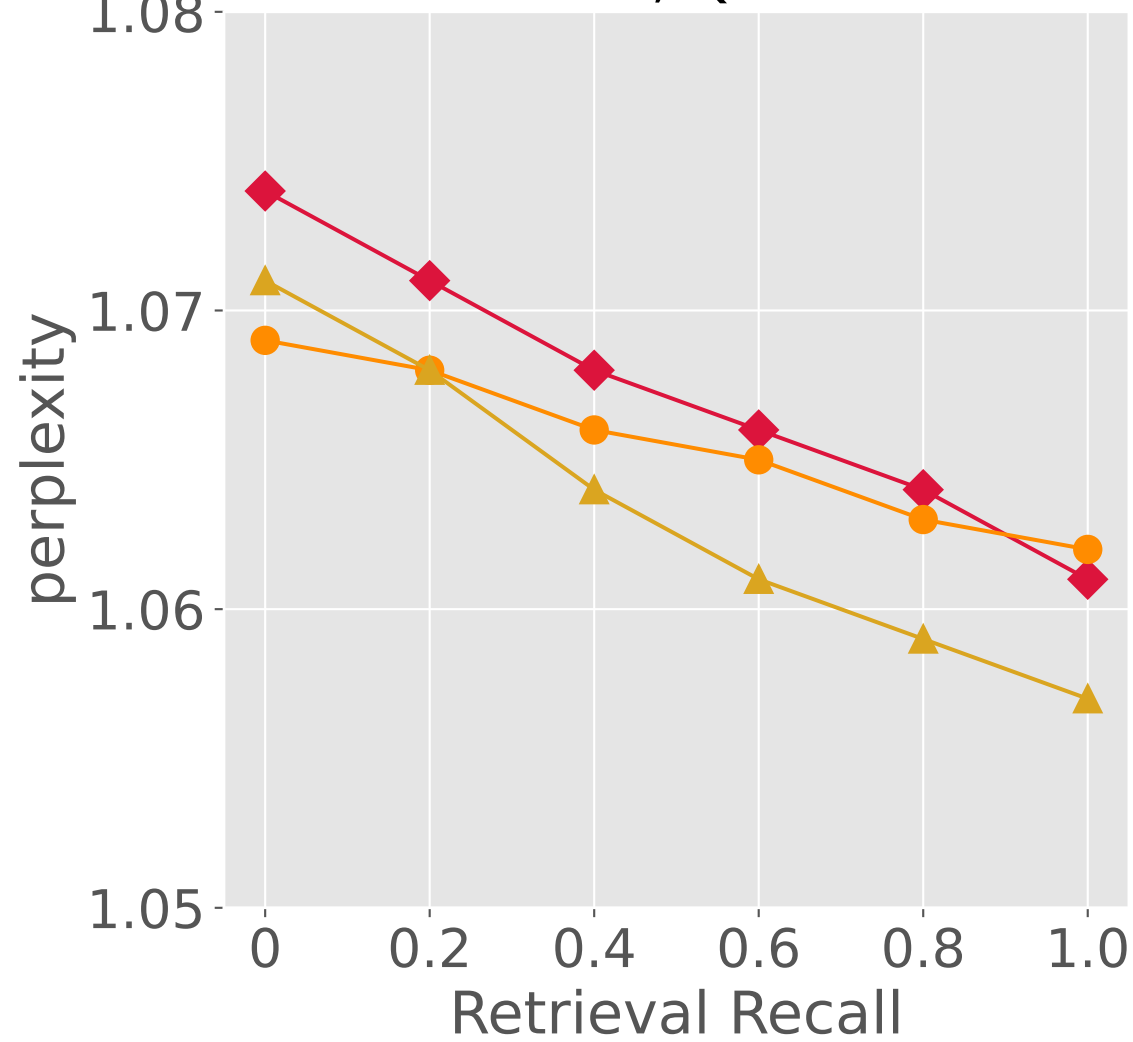
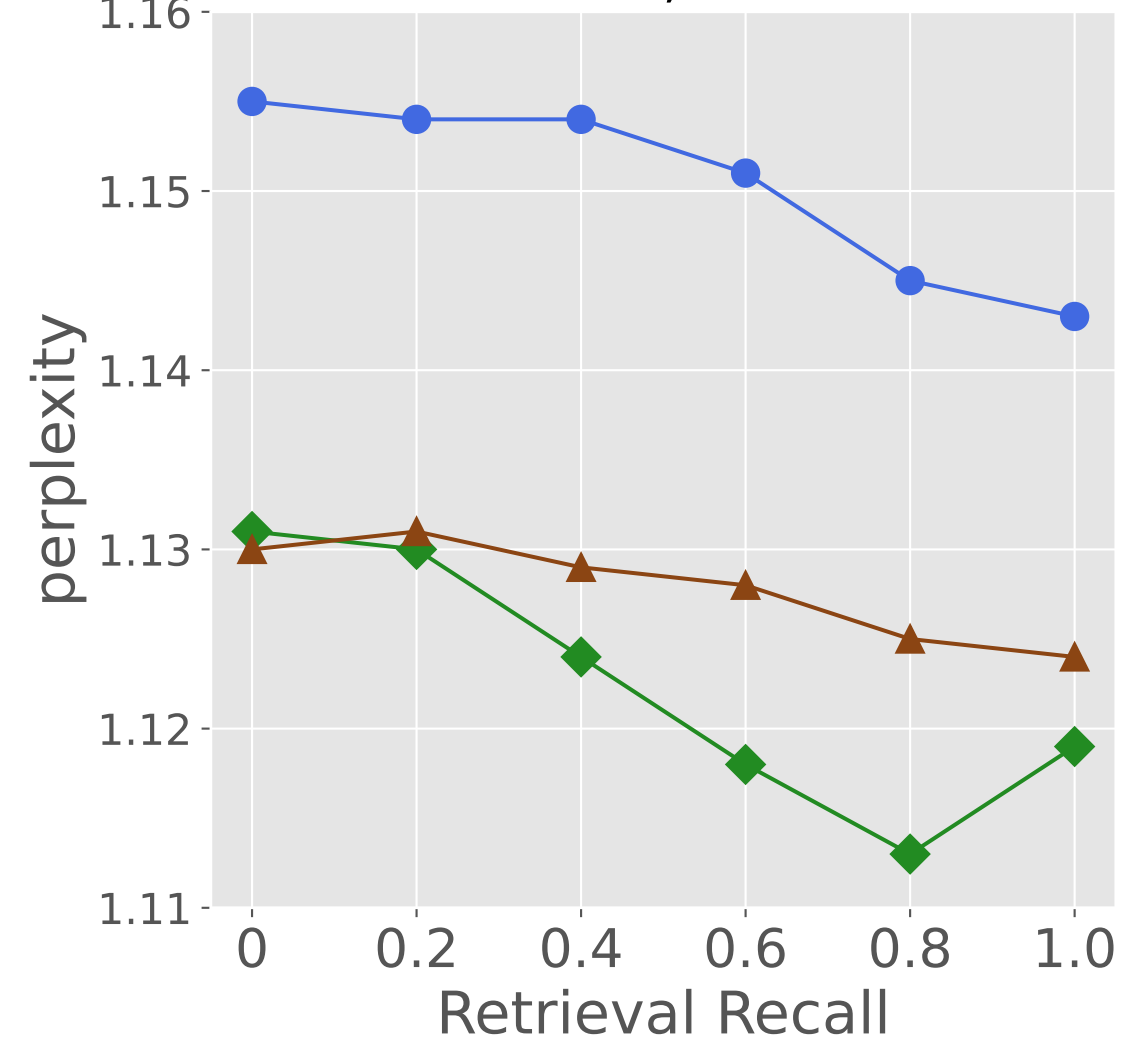


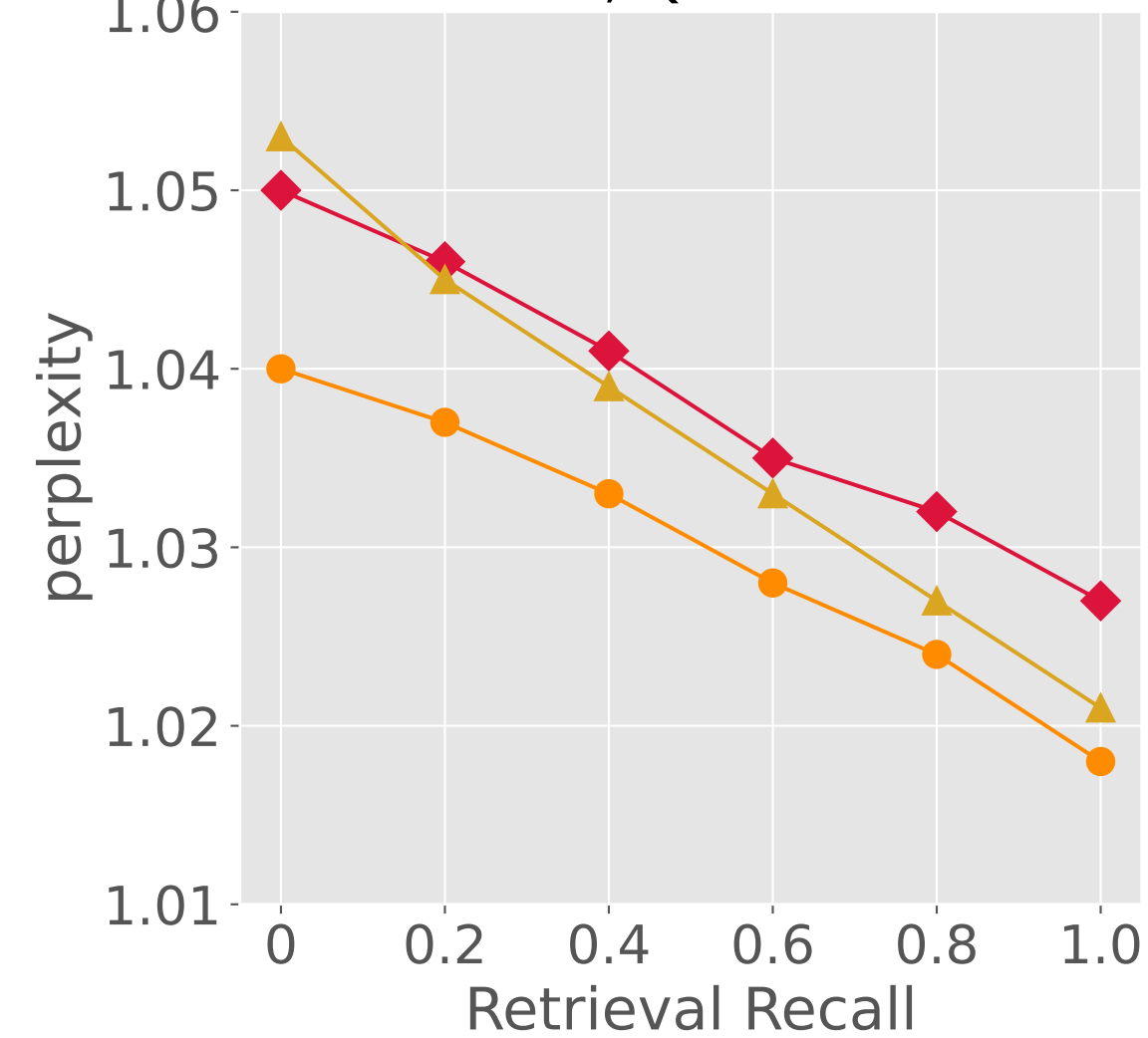
Llama2-13B, QA Datasets



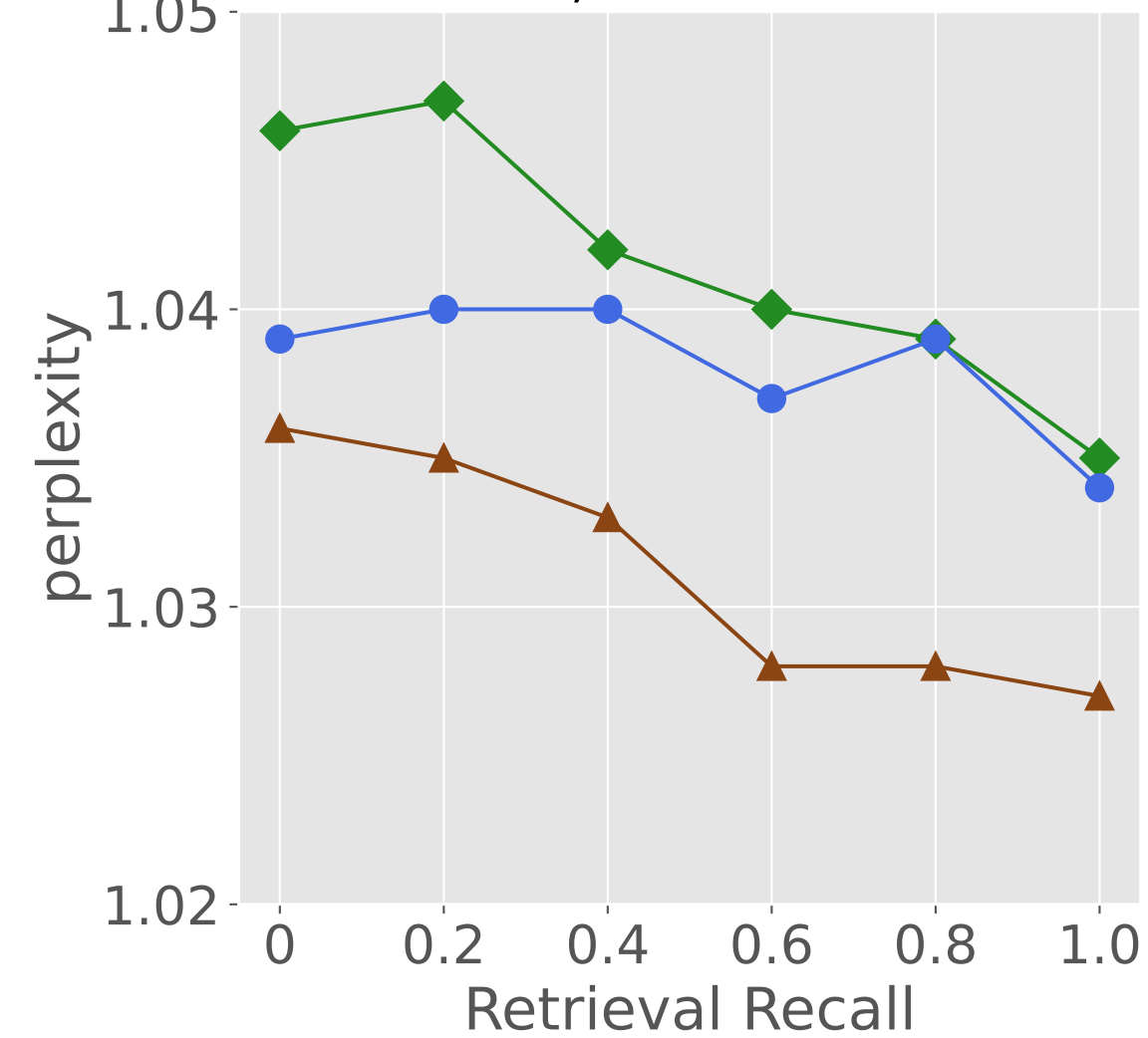
CodeLlama-13B, Code Datasets



GPT-3.5, QA Datasets



GPT-3.5, Code Datasets



Legend: NQ (red diamond), TriviaQA (orange circle), HotpotQA (yellow triangle), CoNaLa (green diamond), DS1000 (blue circle), PNE (brown triangle)