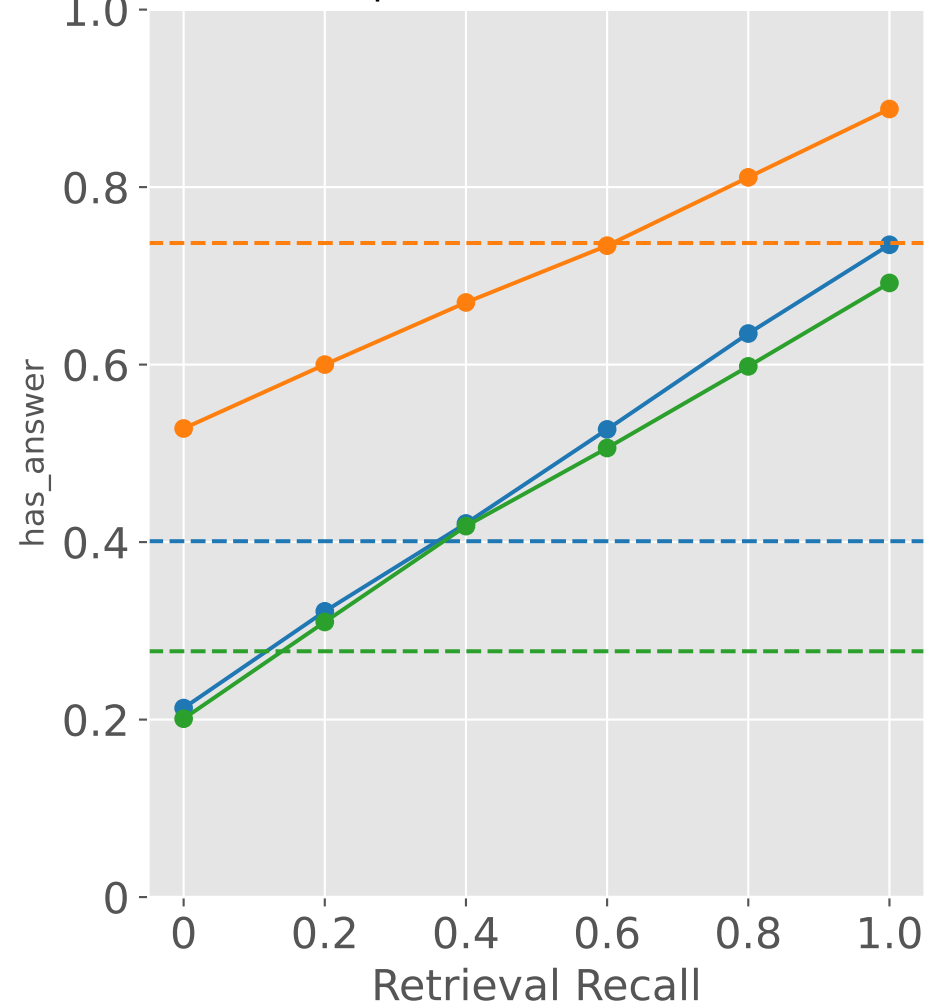
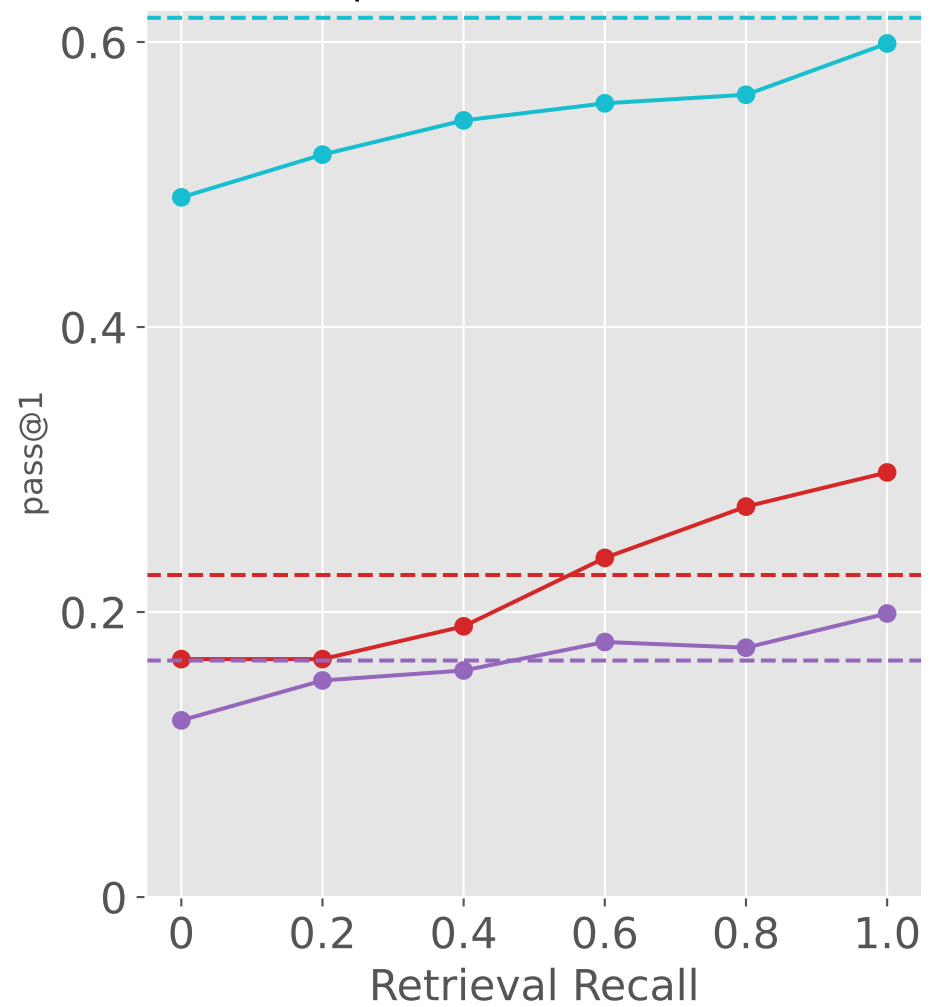


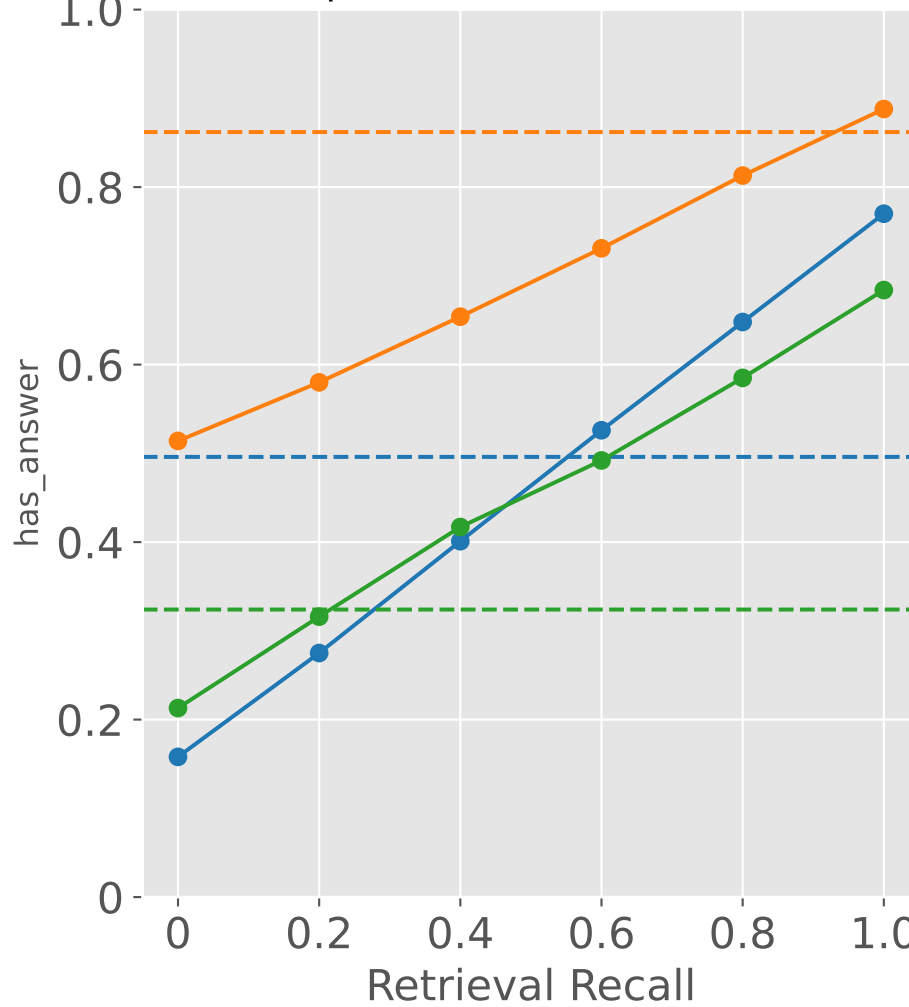
Llama2-13B performances on QA datasets



Llama2-13B performances on code datasets



GPT-3.5 performances on QA datasets



GPT-3.5 performances on code datasets

