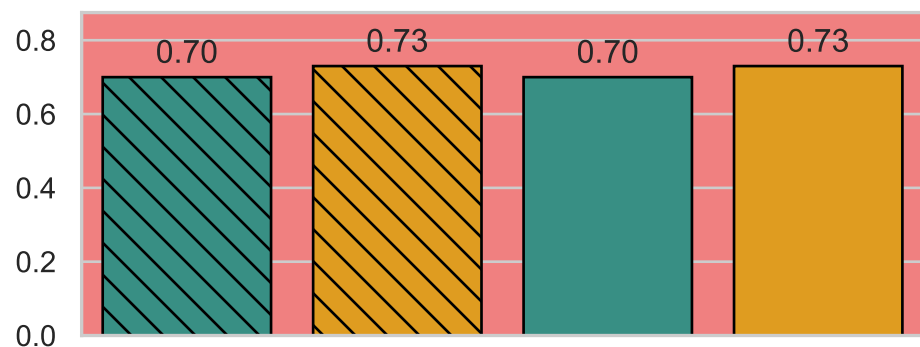
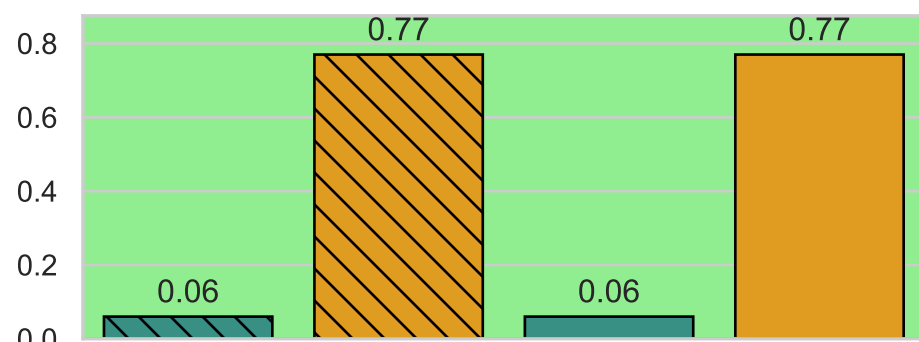


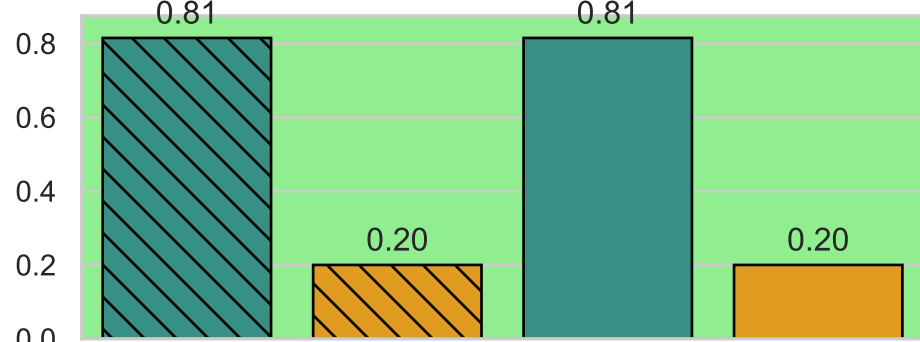
HotpotQA



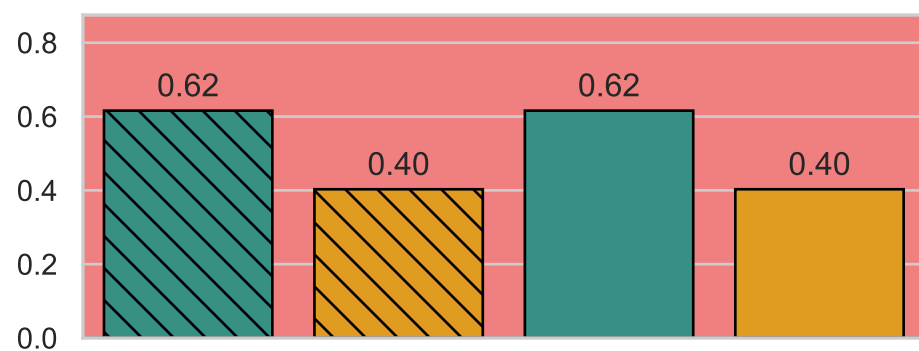
MultispanQA



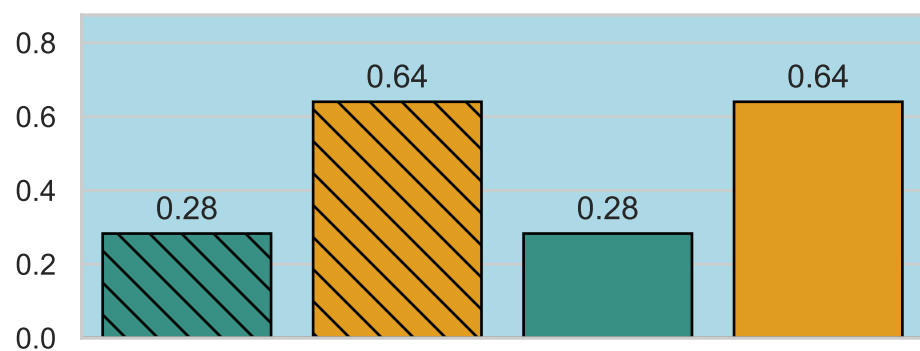
24 Game



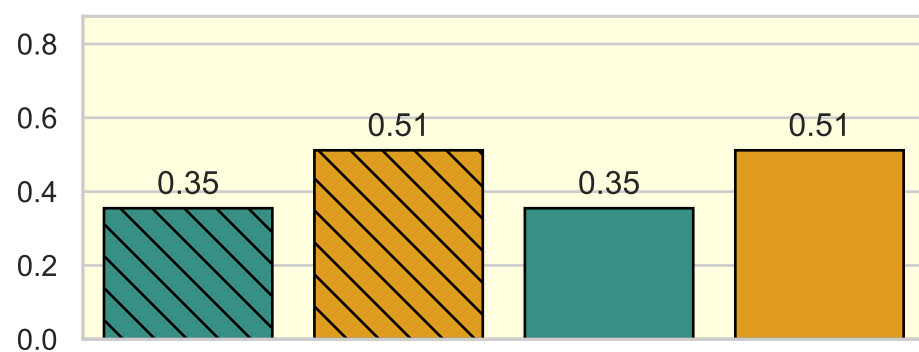
GSM8K



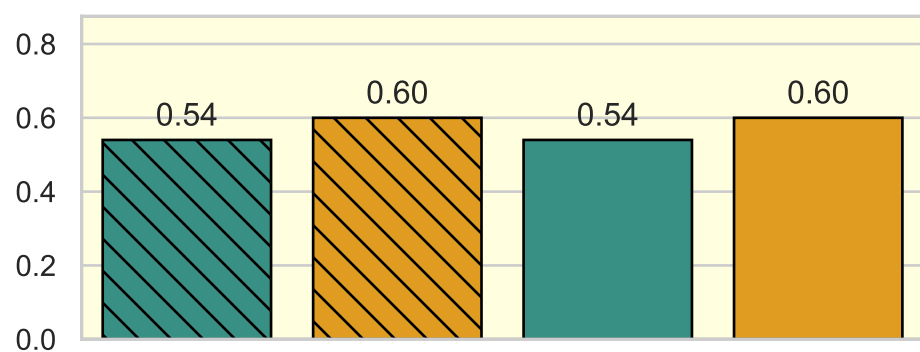
MMLU



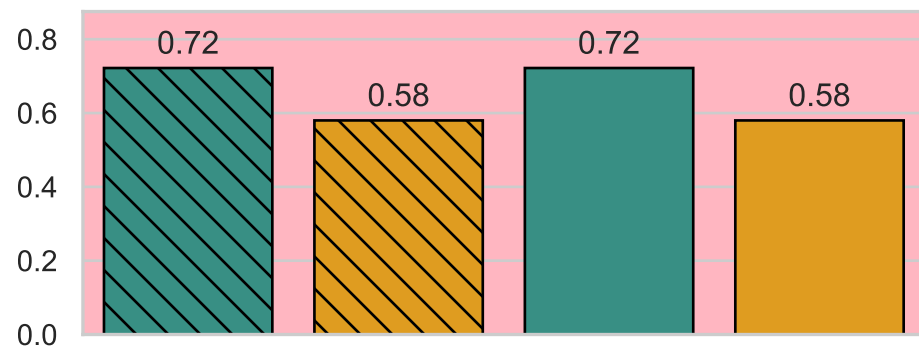
Collie



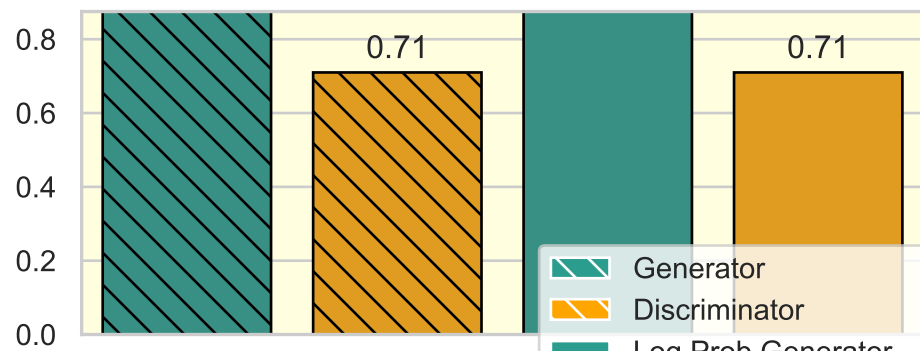
CSQA



Hellaswag



RACE



PIQA

