Abstain-QA

p_{fail}

p_{succeed}

GPT-3.5

GPT-4 32K