

Freq. of Primitives in Attacks with ASR>0.8
bandit_self_score_mixed_meta-llama_Meta-Llama-3-8B-Instruct

