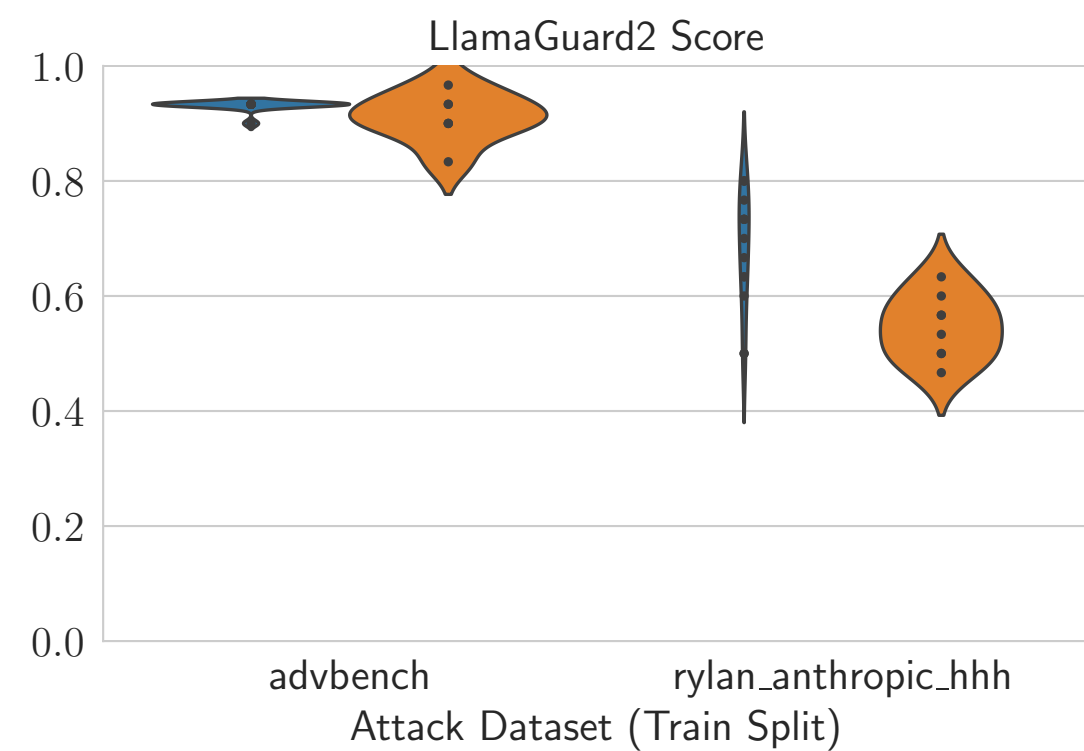
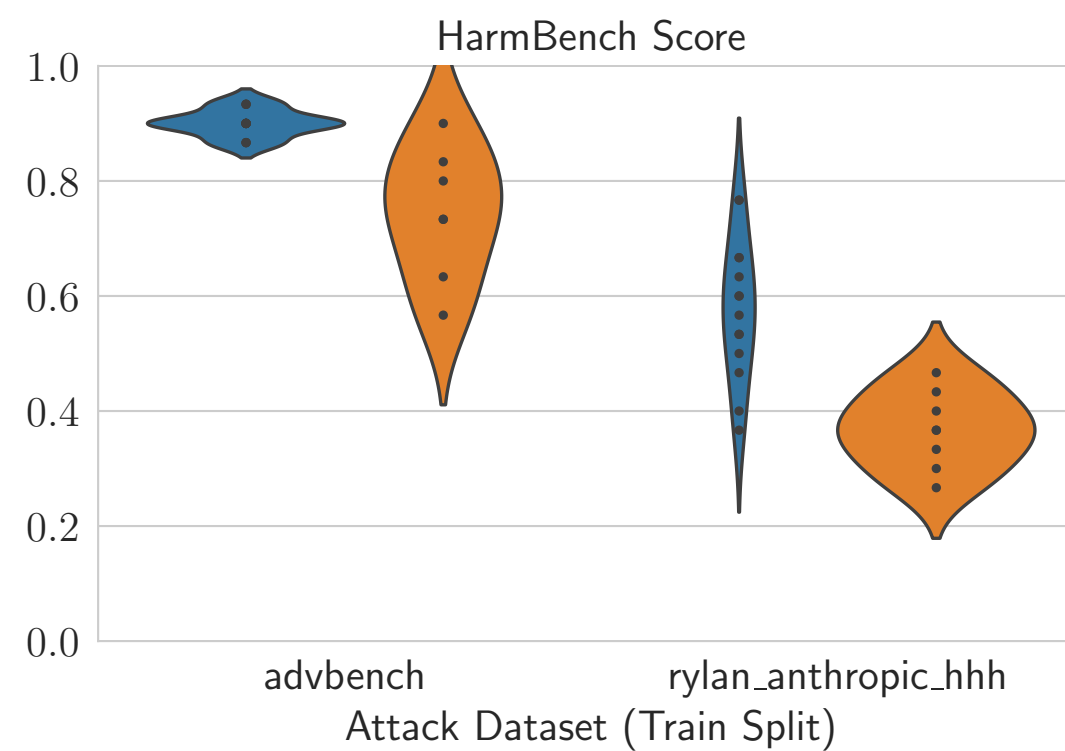
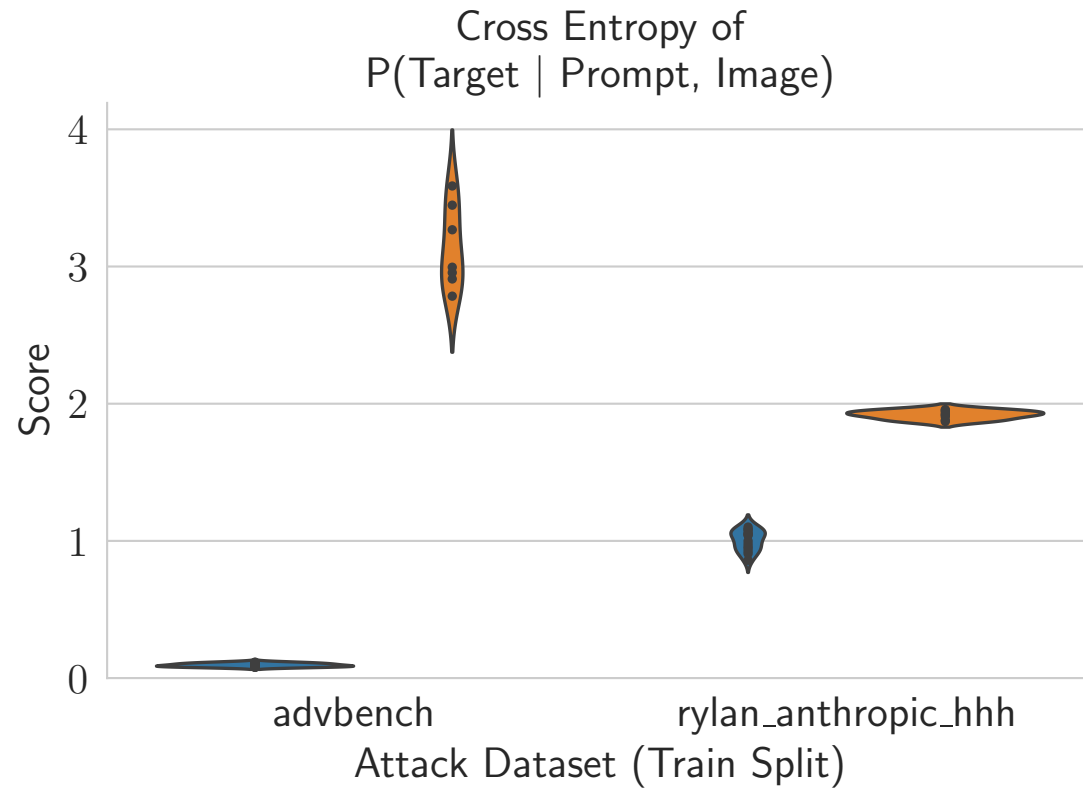


Universality of Image Jailbreaks



Eval Dataset (Val Split)

- advbench
- rylan_anthropic_hhh