

Probe weight cosine (steering\_detectable\_format\_json\_format) steer\_18 probe26

