

Probe weight cosine (steering\_detectable\_format\_json\_format) steer\_0 probe1

