

Probe weight cosine (steering_detectable_format_json_format) steer_8 probe8

Weight index

