

Probe weight cosine (steering_detectable_format_json_format) steer_9 probe9

Weight index

