

Normalized AlpacaEval 2.0 win rate

