

epsilon-greedy with recency-weighted reward average, nonstationary bandit

