

Bandit Assignment Report

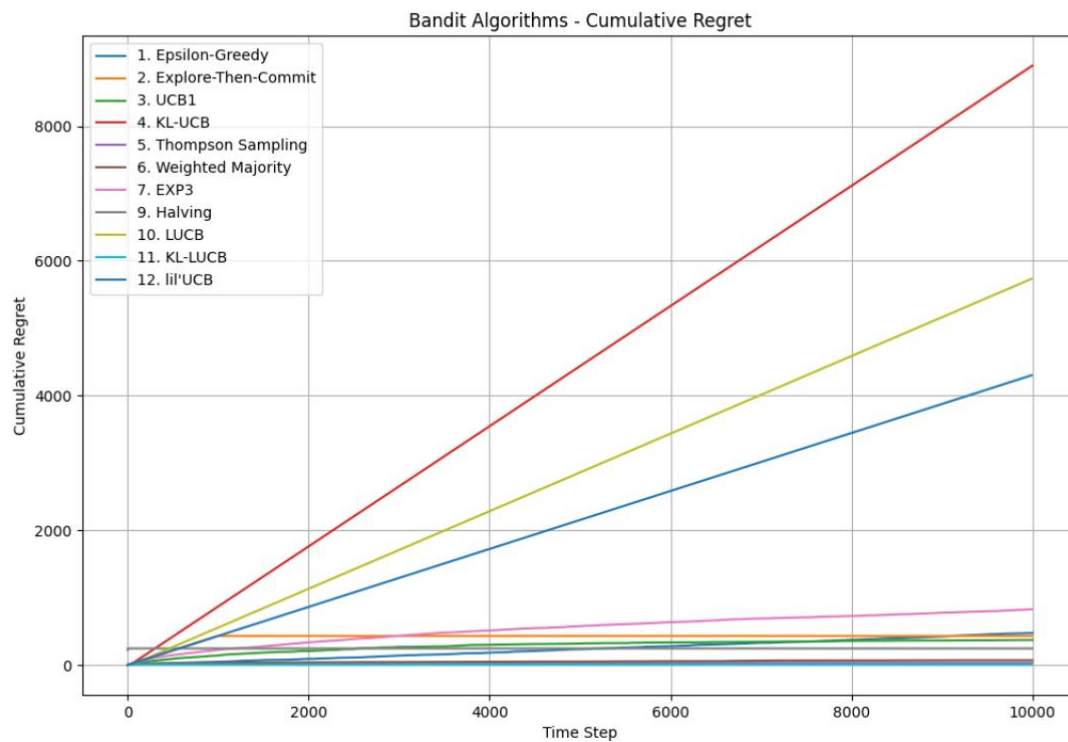
Brief Description of Each Algorithm

1. **Epsilon-Greedy** Chooses the best known arm most of the time, but explores randomly with probability epsilon.
2. **Explore-Then-Commit** Explores each arm a fixed number of times, then commits to the best for the rest.
3. **UCB1** Selects arms based on optimistic confidence bounds derived from empirical mean and uncertainty.
4. **KL-UCB** Like UCB1 but uses KL divergence to compute tighter confidence bounds for Bernoulli rewards.
5. **Thompson Sampling** Bayesian approach that samples from posterior distributions to guide exploration.
6. **Weighted Majority** Adversarial setting algorithm that adjusts arm weights based on estimated losses.
7. **EXP3** Balances exploration/exploitation using exponential-weighted updates in adversarial environments.
8. **LinUCB** Contextual bandit algorithm that models reward as a linear function of features per arm.
9. **Halving** Eliminates half the arms in each round based on empirical means to identify best arm.
10. **LUCB** Compares confidence intervals of top arms and prunes those with overlapping intervals.
11. **KL-LUCB** Variant of LUCB using KL-divergence-based confidence bounds for tighter selection.
12. **lil'UCB** Uses law-of-iterated-logarithm based bounds to optimistically guide exploration.

Implementation Approach

All 12 bandit algorithms are implemented from scratch in Python using function-based style. The simulation environment supports both Bernoulli and Gaussian rewards. The contextual bandit setup generates synthetic d-dimensional features and uses linear reward models. Evaluation includes cumulative regret, best-arm frequency, and confidence bounds.

Experimental Plots & Analysis



Algorithm	Suitable For	Type
Epsilon-Greedy	Stochastic rewards	Basic
Explore-Then-Commit	Stochastic rewards	Basic
UCB1	Stochastic rewards	Basic
KL-UCB	Bernoulli rewards	Basic
Thompson Sampling	Stochastic rewards	Bayesian
Weighted Majority	Adversarial rewards	Adversarial
EXP3	Adversarial rewards	Adversarial
LinUCB	Contextual rewards	Contextual
Halving	Best arm ID	Pure exploration
LUCB	Best arm ID	Pure exploration
KL-LUCB	Best arm ID	Pure exploration
lil'UCB	Best arm ID	Pure exploration