

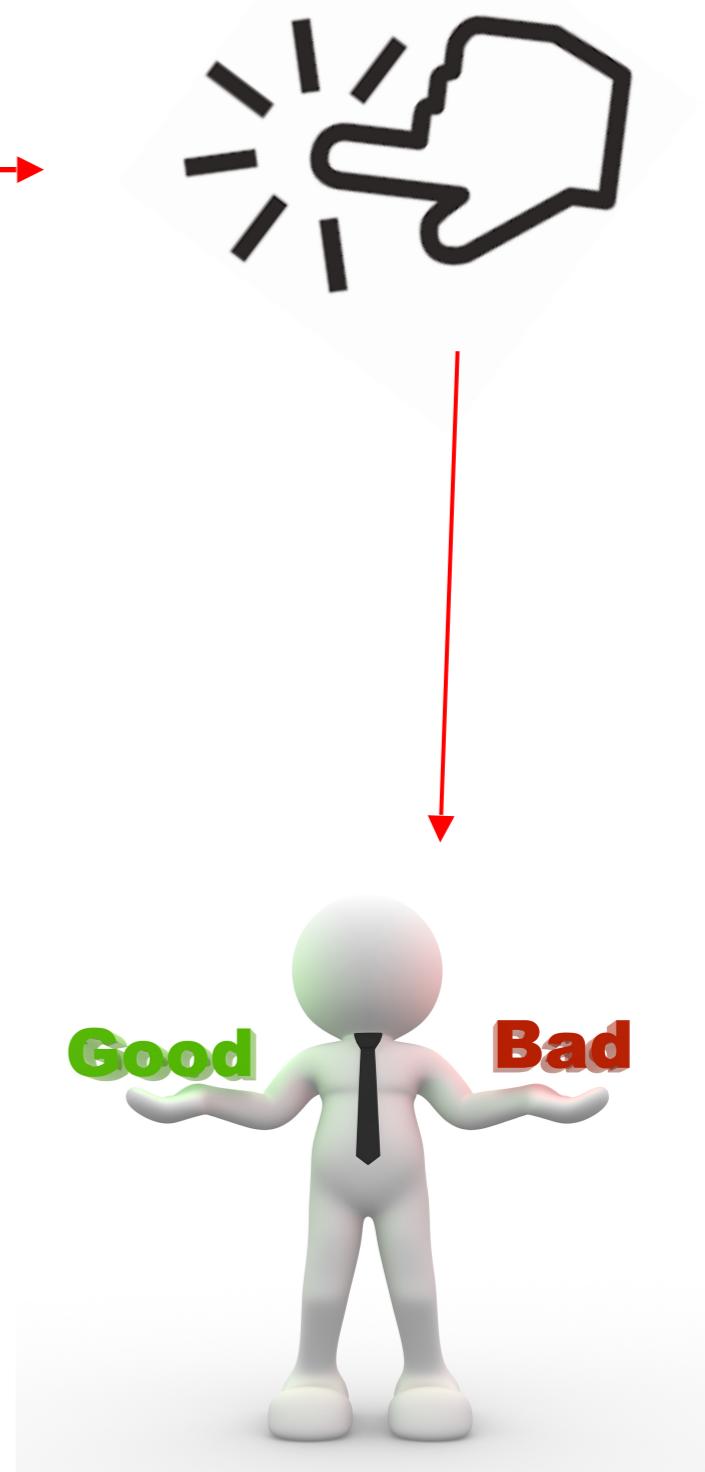
A Bandit Approach to Personalized News Article Recommendation



Qingyun Wu

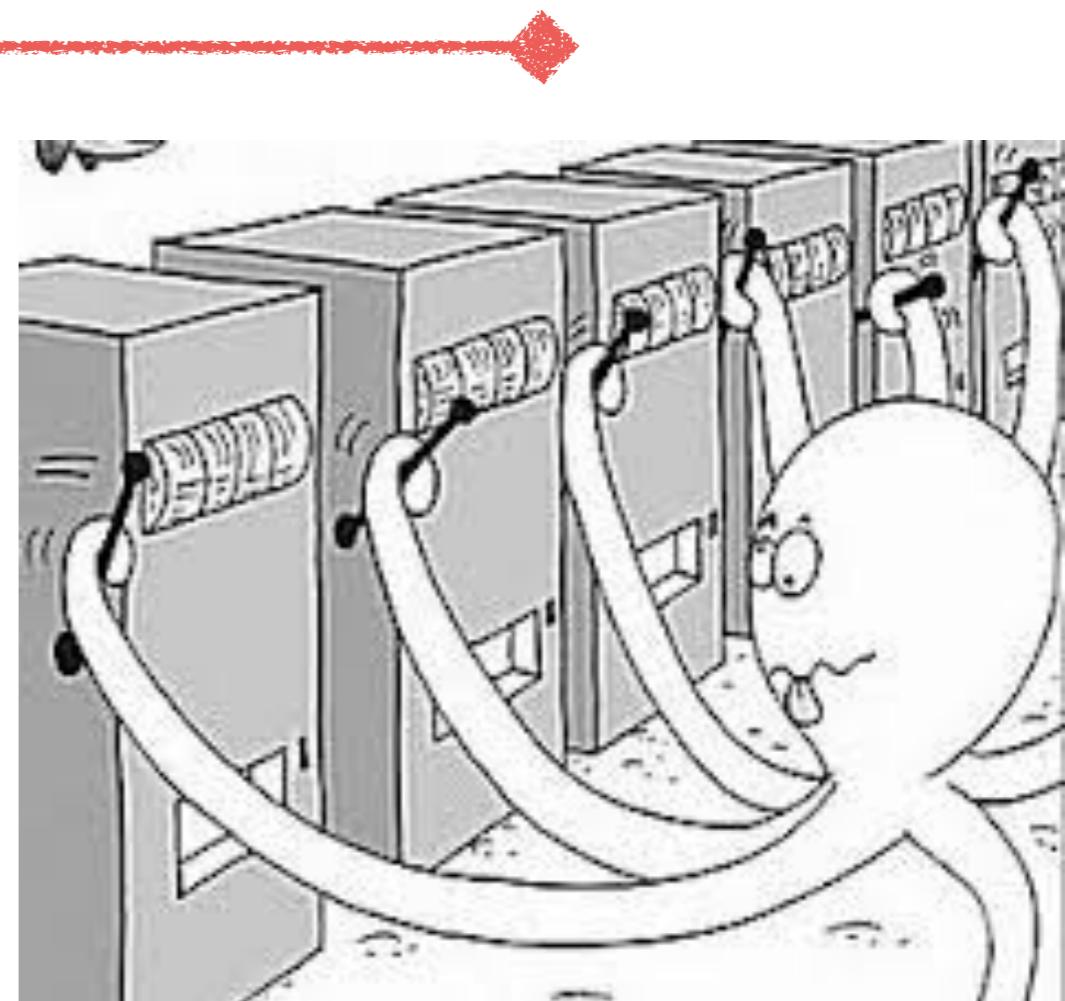
News Recommendation Cycle

The screenshot shows the Yahoo News homepage. On the left sidebar, there are links for 'Photos' and 'Recommended Games' (BINGO, SUPER HERO SLOTS). The main content area features a news article titled 'Dayton reaches first Sweet 16 with 99-94 upset of Kentucky'. Below it is another article, 'Guess Who's About To Go Bankrupt in America', followed by 'Chadian helicopters bomb Boko Haram on Nigeria-Niger border: Niger army', 'Death of Lee Kuan Yew: Live Report', and 'Human remains found in Alaska, may be tied to missing family'. At the top right of the page are 'Search News' and 'Search Web' buttons.



A K-armed Bandit Formulation

- A **gambler** must decide which of the K non-identical slot machines(we called them **arms**) to **play** in a sequence of trials in order to maximize total reward.



News Website \longleftrightarrow gambler
Candidate news articles \longleftrightarrow arms
User Click \longleftrightarrow Reward

How to pull arms to maximize reward?

How to select articles to serve users to maximize user clicks

Ideal Solution



Pick $\arg \max_a \mu_a$

But we DO NOT know the mean.



Let's estimate it

Choices	X_1	X_2	X_3	X_4	X_5	X_6	...
a_1					1	1	
a_2	0		1	0			
...							
a_k		0					

Time →



Exploitation VS. Exploration

Exploitation: pull an arm for which we current have the highest estimate of mean of reward



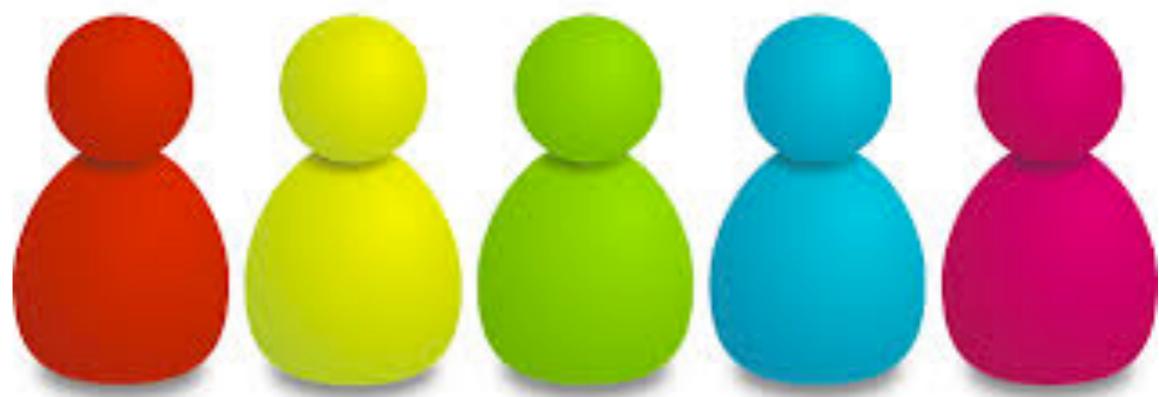
Exploration: Pull an arm we never pulled before

Not only look at the **mean**, but also the **confidence!**

Pick $\hat{\arg \max}_a (\mu_a + \alpha * UCB)$

UCB1

$$\arg \max_a (\hat{\mu}_a + \sqrt{\frac{2 \ln T}{n_a}})$$



LinUCB (Contextual)

- **Article feature:** URL categories, topic categories
- **User feature:** demographic information, geographic features, behavioral categories

LinUCB(Contextual)

Assumption

$$E(y_{t,n} | x_{t,n}) = x_{t,n}^T \theta_n$$

Article Feature Vector

User preference

Parameter Estimation

$$\hat{\theta}_n = A^{-1}b$$
$$A_n = \lambda I + \sum_t x_{t,n} x_{t,n}^T \quad b_n = \sum_t y_{t,n} x_{t,n}$$

Pick $\arg \max_a (x_{t,n}^T \hat{\theta}_n + \alpha \sqrt{x_{t,n}^T (D_n^T D_n + I_d) x_{t,a}})$



From LinUCB to Collaborative-LinUCB



$$\sum_{i=1}^N W_{ij} = 1 \quad \sum_{j=1}^N W_{ij} = 1$$

If user i and user j are connected
by an edge, $W_{ij} > 0$

Otherwise $W_{ij} = 0$

Assumption

$$E(r_{t,n} | x_{t,n}) = x_{t,n}^T \theta_n$$



$$E(r_{t,n} | x_{t,n}) = x_{t,n}^T \sum_j W_{nj} \theta_j$$

Collaborative-LinUCB



Parameter Estimation

$$\hat{\theta}_n = A_n^{-1} b_n$$

$$A_n = \lambda I + \sum_{m=1}^N W_{mn}^2 \sum_t x_{tm} x_{tm}^T$$

$$b_n = \sum_{m=1}^N W_{mn} \sum_t (x_{tm} y_{tm} - x_{tm} x_{tm}^T \sum_{j \neq n} W_{mj} \theta_j^U)$$

Make a choice

$$\arg \max_a \left(x_{tn}^T \sum_{j=1}^N \hat{\theta}_{nj} + \alpha \sqrt{x_{tn}^T \sum_{j=1}^N W_{nj} A_j^{-1} x_{tn}} \right)$$

Performance Evaluation



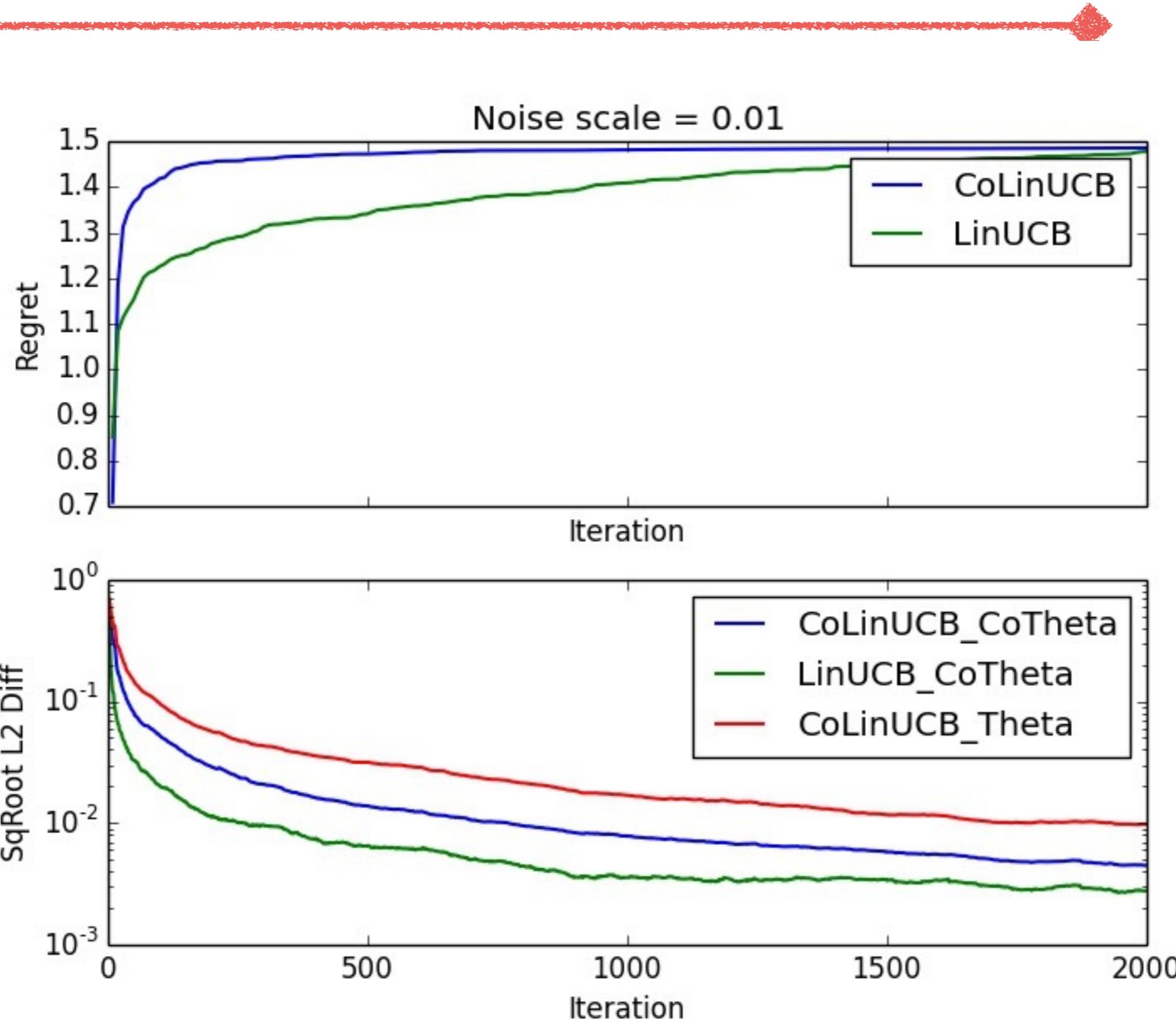
Measurement criteria

Regret

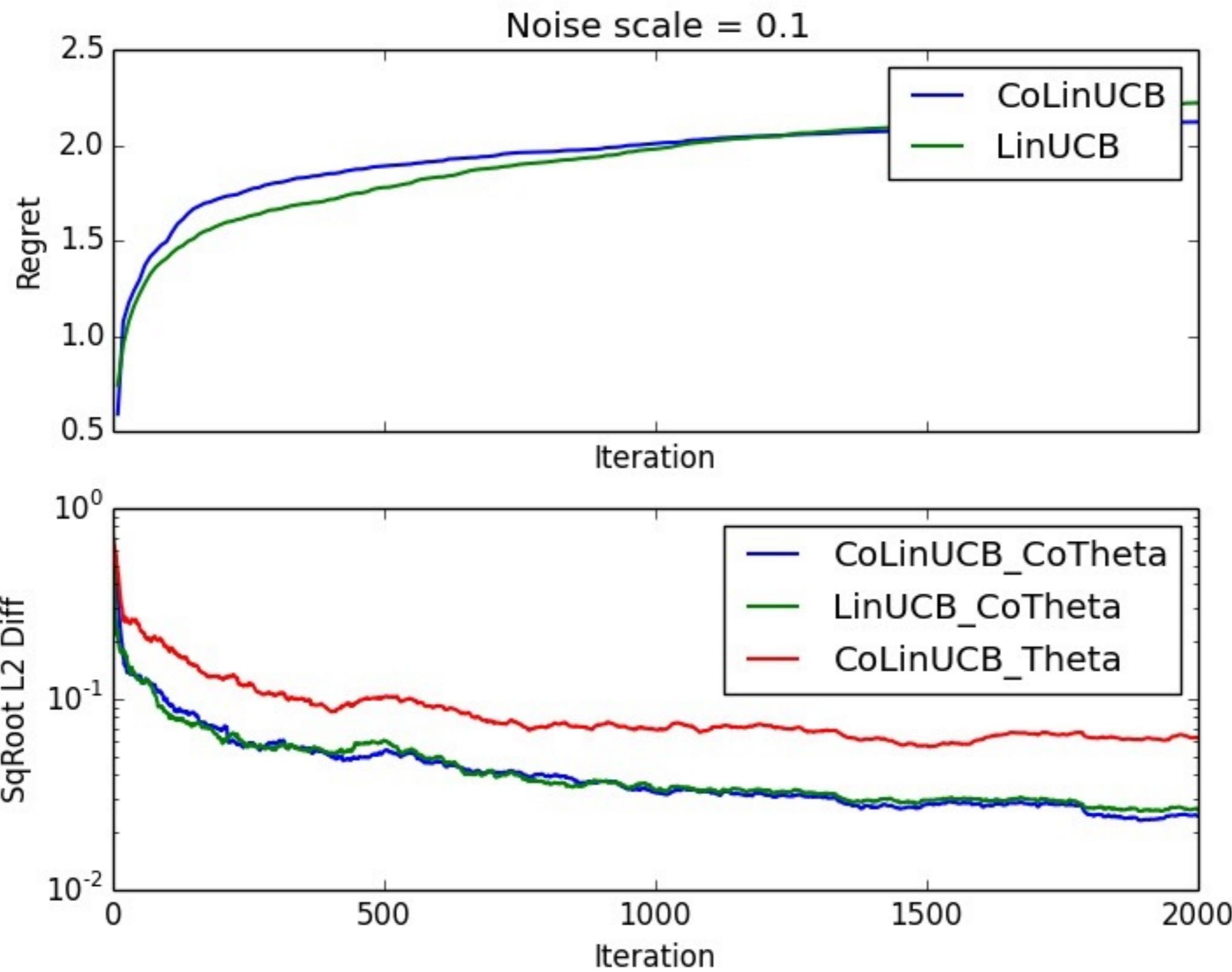


$$R_A(T) = E\left[\sum_t r_{t,a_t^*}\right] - E\left[\sum_t r_{t,a_t}\right]$$

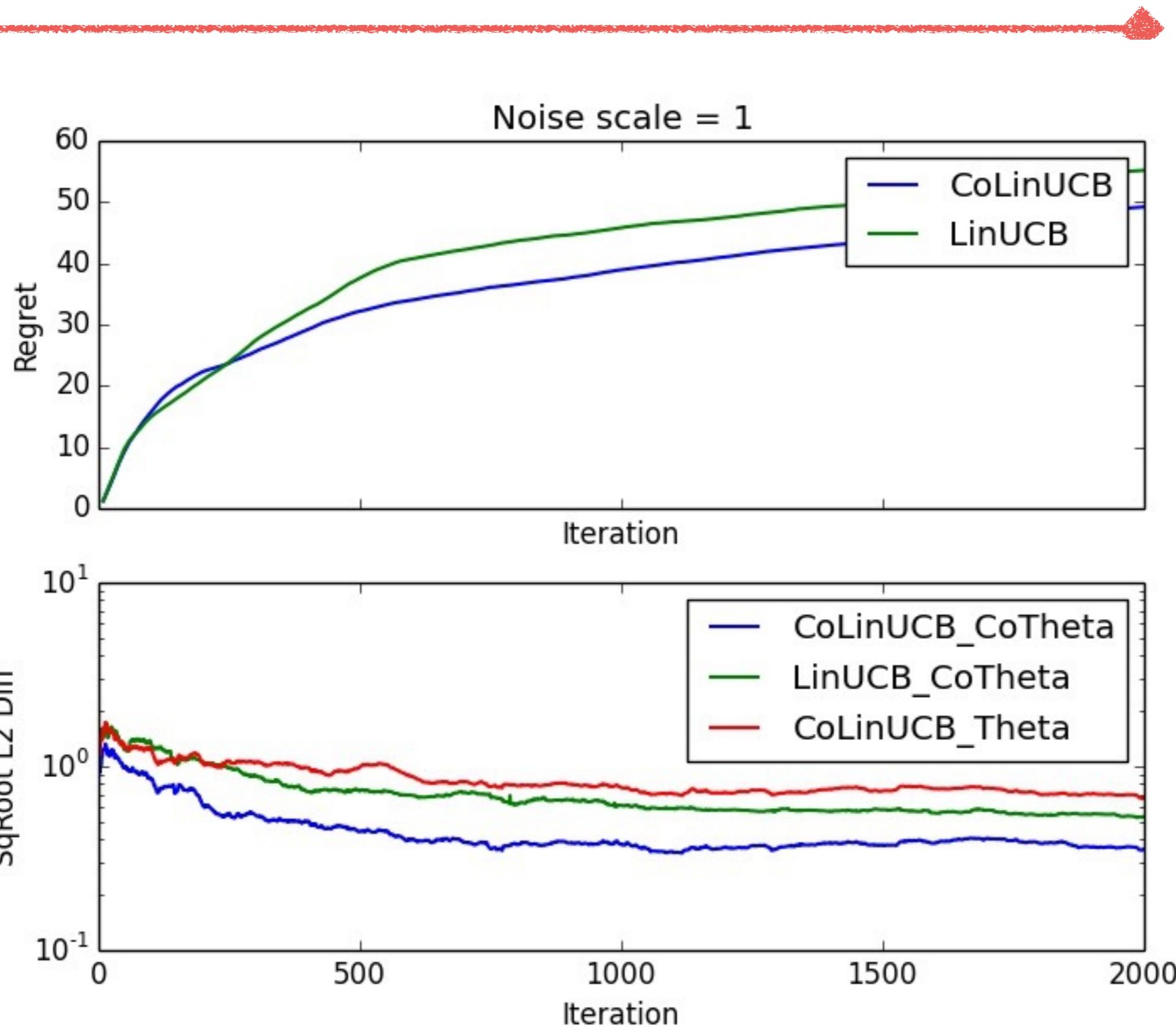
Performance Evaluation



Performance Evaluation



Performance Evaluation



Summary



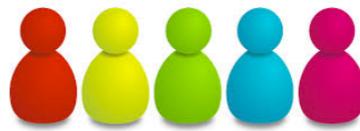
CoLinUCB



UCB1



LinUCB



Q&A

