BT4014 Analytics Driven Design of Adaptive Systems

Group Project, Question 2

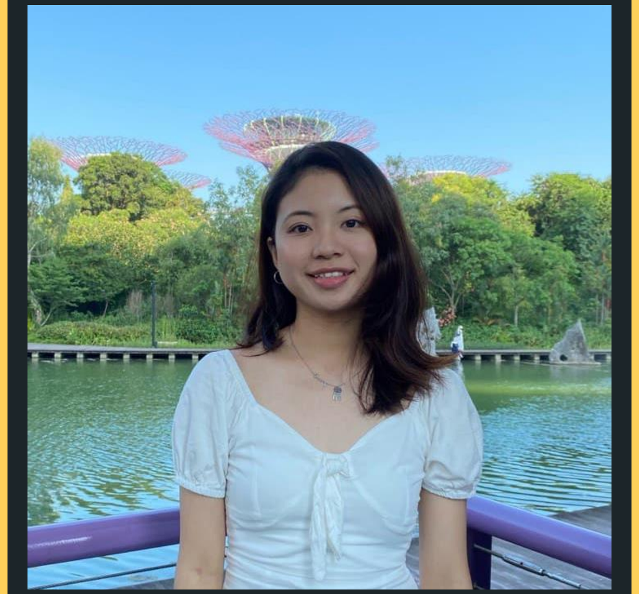
BUILDING A BEER RECOMMENDER

with the use of Bandit Algorithms

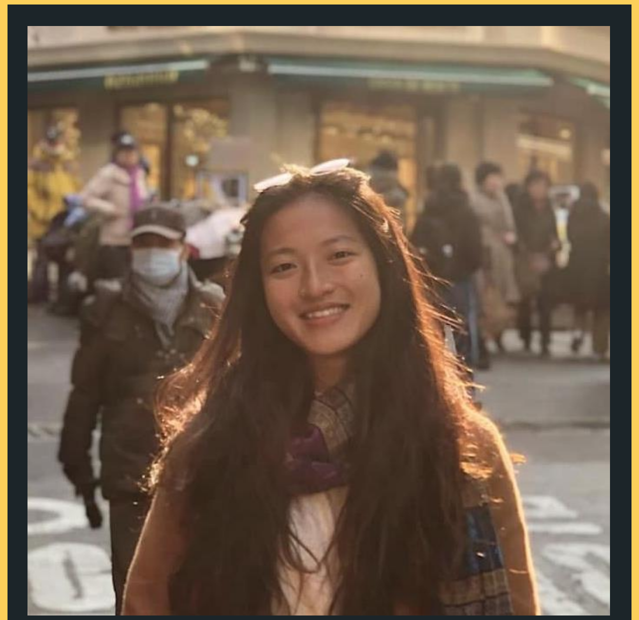
Prepared by: Group 7
Chan Cheah Cha, Chua Kai Bing, Goh Jia Yi, Lim Jia Qi, Tan Zen Wei
A0189006A, A0185606Y, A0185610J, A0189626M, A0188424X

MEET THE TEAM

TAN **ZEN WEI**



CHUA **KAI BING**



CHAN **CHEAH** CHA

LIM **JIA QI**

GOH **JIA YI**, JESA

AGENDA

1. Introduction
2. Data Exploration & Preprocessing
3. Collaborative Filtering
4. Multi-Armed Bandit
5. Algorithm Selection
6. Future Improvements
7. Conclusion

1. INTRODUCTION

BEER INDUSTRY

Global Market Value expected to rise to $0.685 billion by 2025

> 19,000 brewing companies

> 300,000 beers listed

BRAND NAMES

Consumers tend to stick with well-established brands

→ Harder for smaller companies to stand out

COVID-19

Shift in consumer behaviour to sample new brands

→ Growth in smaller companies

OUR STANCE

LICENSED INTERMEDIARY IN BEER DISTRIBUTION

Build a recommender system on our eCommerce platform to recommend less commonly found but still higher quality beers for consumers

BENEFITS

Consumers

Ease their efforts in exploring new brands

Beer Companies

Help smaller companies to stand out from big brands

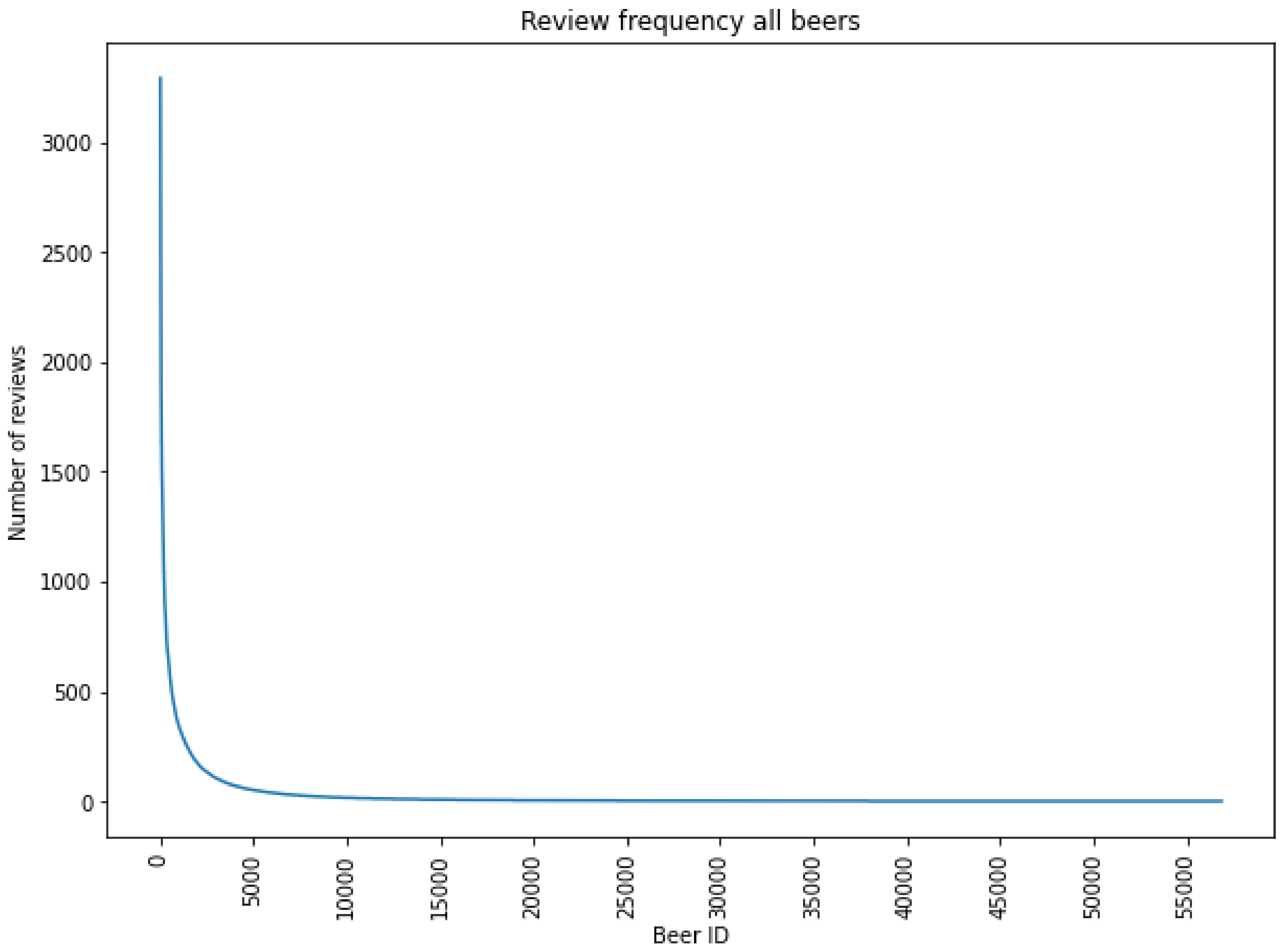
Our Business

Better inventory management,
Increase consumer engagement

2. DATA EXPLORATION & PREPROCESSING

BEER REVIEWS DATASET

- Data collected from BeerAdvocate, hosted on data.world
- ~1.5 million reviews of beers collected over 10 years, up to Nov’11
- 33,387 customers
- 66,055 beer brands
- Overall ratings on a scale from 1 (worst) to 5 (best)



REVIEW FREQUENCY OF ALL BEERS

- Sparse Reviews
- Only ~5% beers received ratings by more than 100 users
- Rest with little or no user interactions

Less predictable for more users & highly sensitive to individuals who are interested in more obscure beers

FILTERING

- Review count
- The more the reviews, the more the popularity

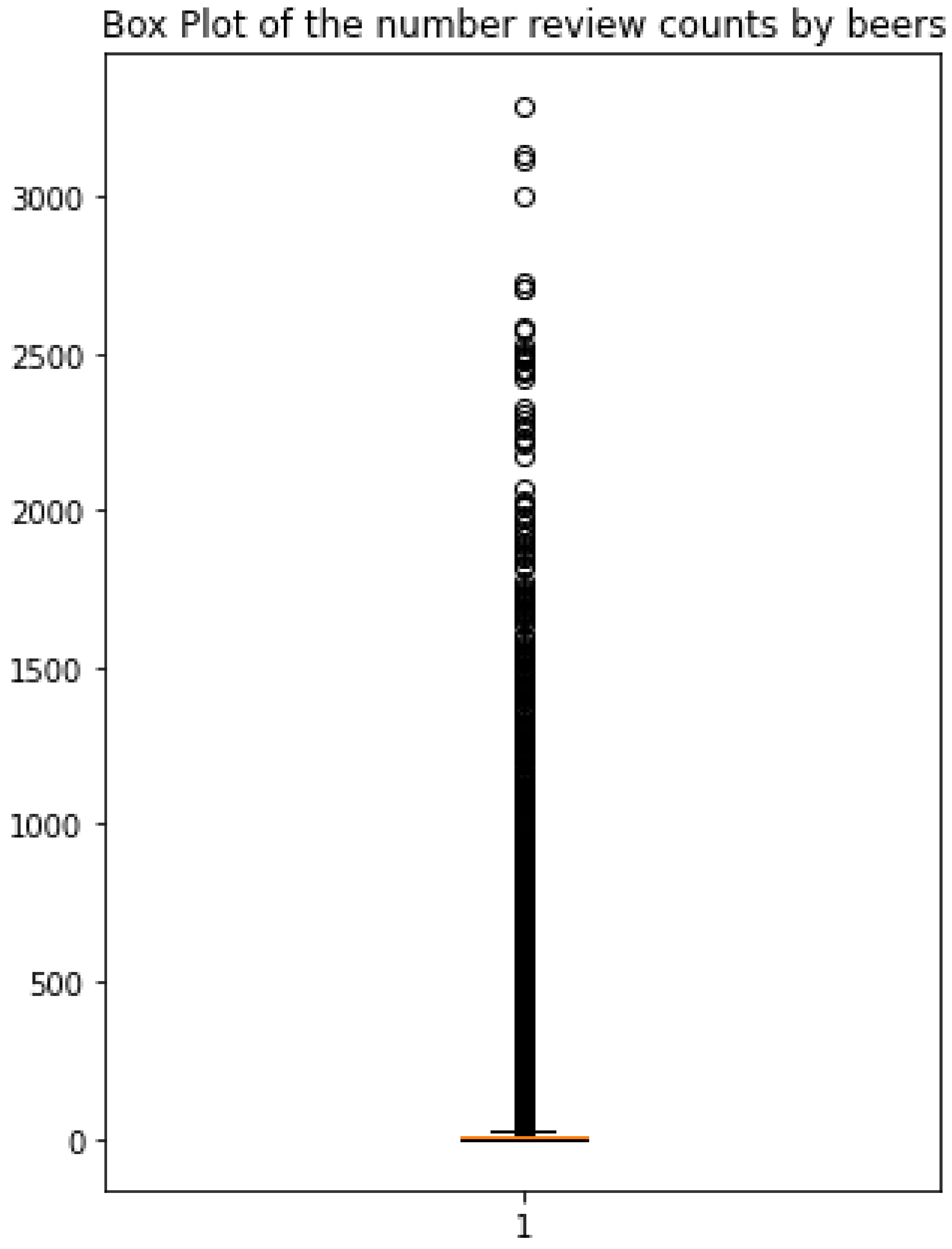
POPULARITY

- Average overall rating
- The higher the average overall rating, the higher the quality

QUALITY

TARGET

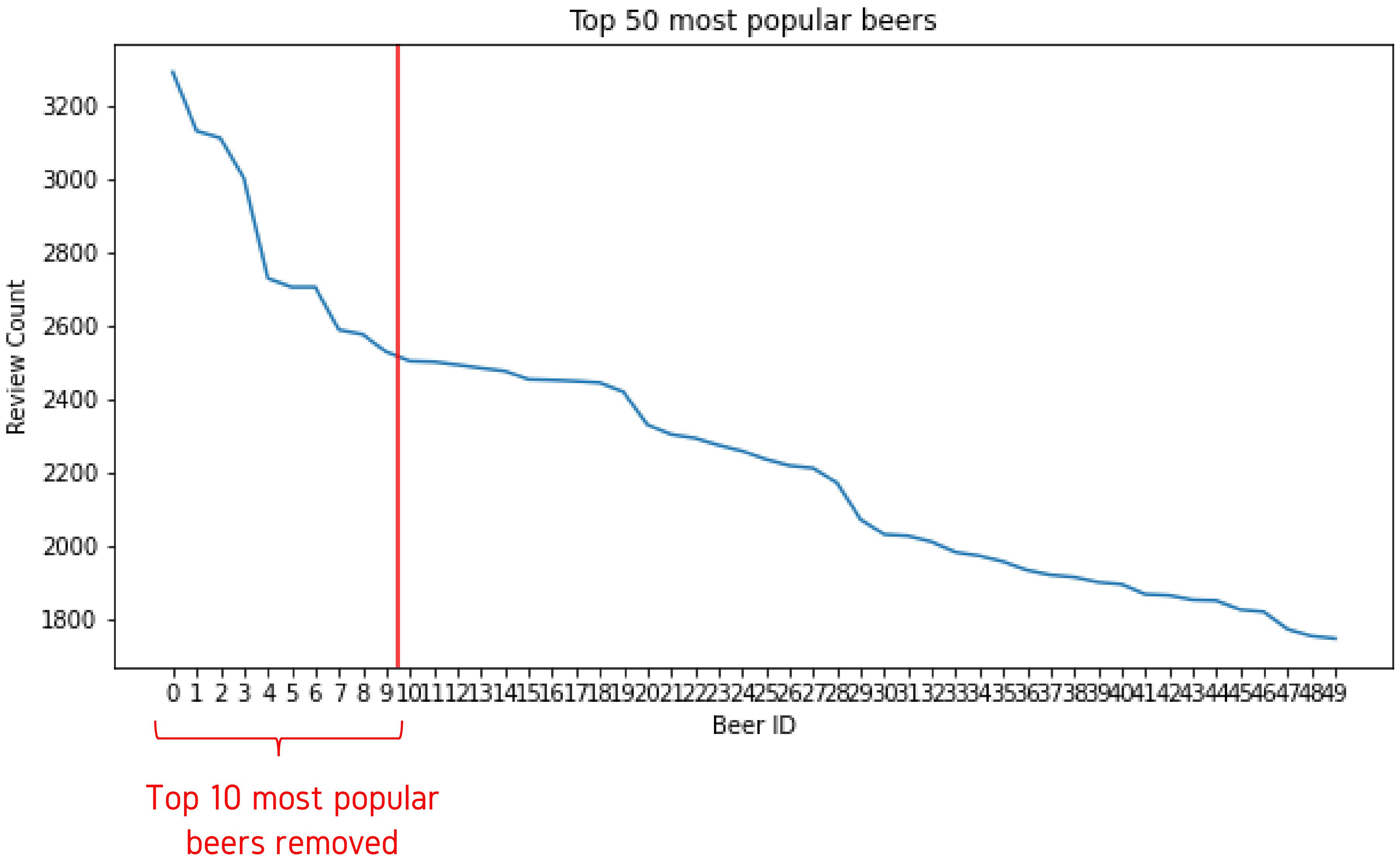
- Less Popular
- Higher Quality



BOXPLOT OF REVIEW COUNT

- Range: 1 to 3,290 reviews
- IQR: 1 to 9 reviews
- Mean: 3 reviews
- Median: 28 reviews

- Distribution highly skewed to the right
- Most with little reviews
- Popular beers are outliers



FILTERING BY POPULARITY

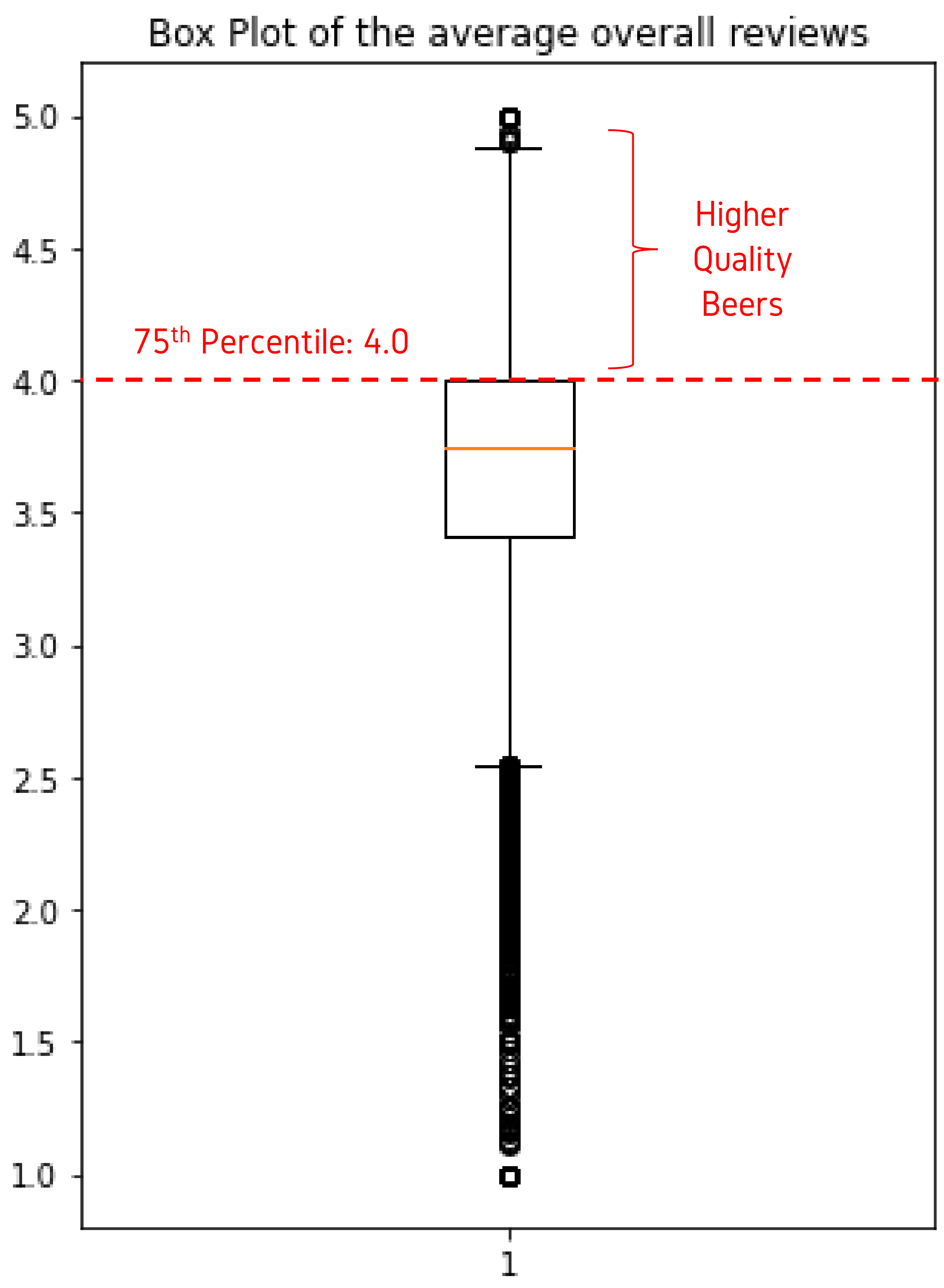
REMOVED TOP 10 BEERS

- Cause of disparity in distribution
- **Assumption:** Review count is an accurate representation of the actual popularity of the beers amongst the population

KEPT BEERS WITH > 100 REVIEWS

- To ensure sufficient reviews and avoid data sparsity issues

56,857 beers reduced to **3,094** beers



FILTERING BY QUALITY

KEPT TOP 100 BEERS

- To find higher quality beers

CHECKING FOR QUALITY

- Top 100 beers have > 4.0 average overall quality (above 75th percentile)

3,094 beers shortlisted to **100** beers

TRANSFORMING DATASET INTO UTILITY MATRIX FOR BANDIT ALGORITHMS

User did not rate beer

User rated the beer

	Citra DIPA	Cantillon Blåbær Lambik	Heady Topper	Deviation - Bottleworks 9th Anniversary	Trappist Westvleteren 12	Pliny The Younger	Founders CBS Imperial Stout	Live Oak HefeWeizen	Portsmouth Kate The Great	Rare Bourbon County Stout	Duck Duck Gooze	Reality Czeck	Weihenstephaner Hefeweissbier	Trappist Westvleteren 8
0	0.0	0.0	0.0	0.0	4.5	5.0	0.0	0.0	0.0	0.0	0.0	0.0	5.0	0.0
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	4.5	0.0	4.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.5
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

- **62,257 reviews** from **10,437 users**
- **100 beers** each representing one column

3. COLLABORATIVE FILTERING

WHAT IS COLLABORATIVE FILTERING?

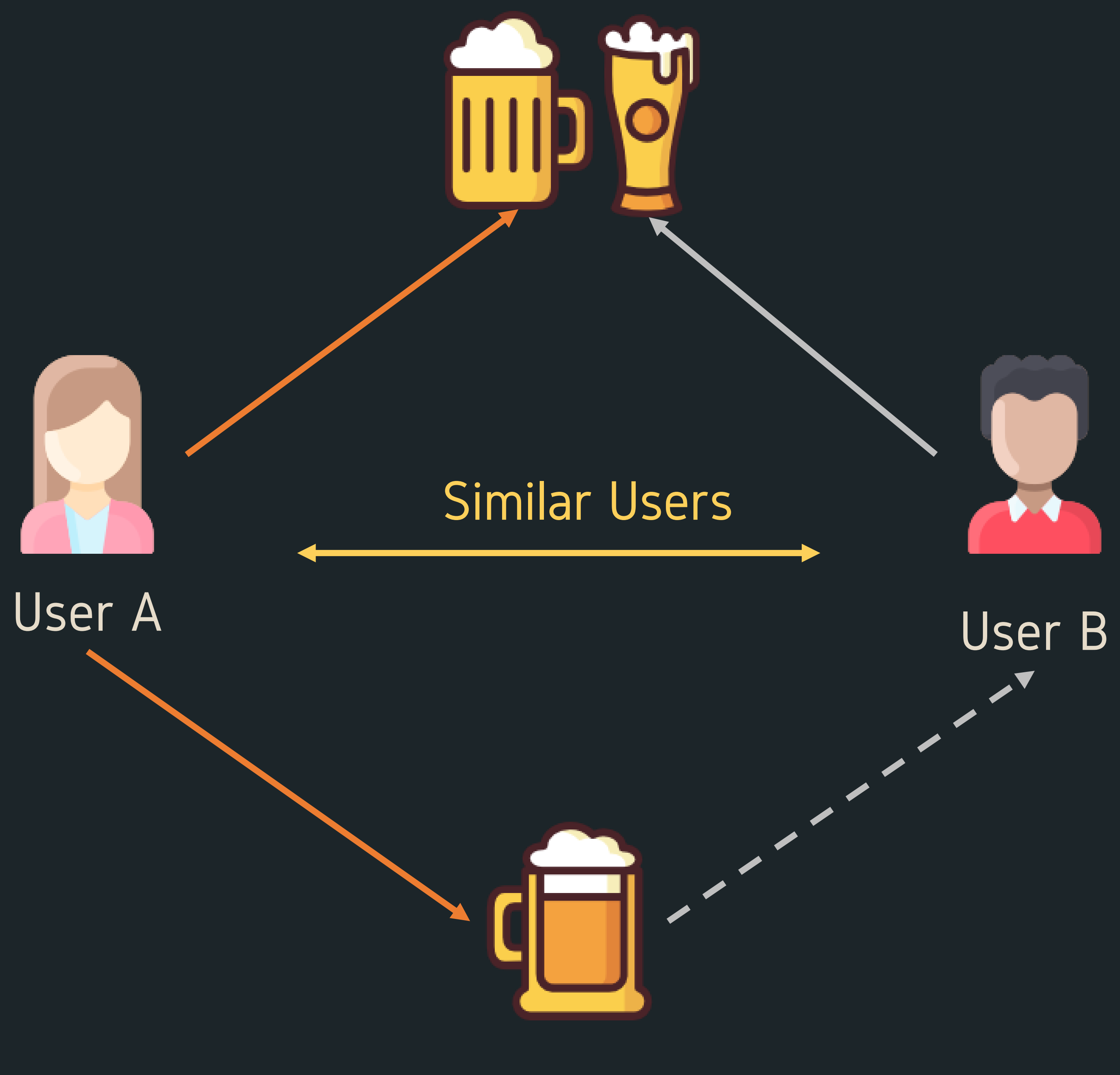
RECOMMENDER SYSTEM

Makes predictions on user preference based on historical data of users' preferences on a set of items

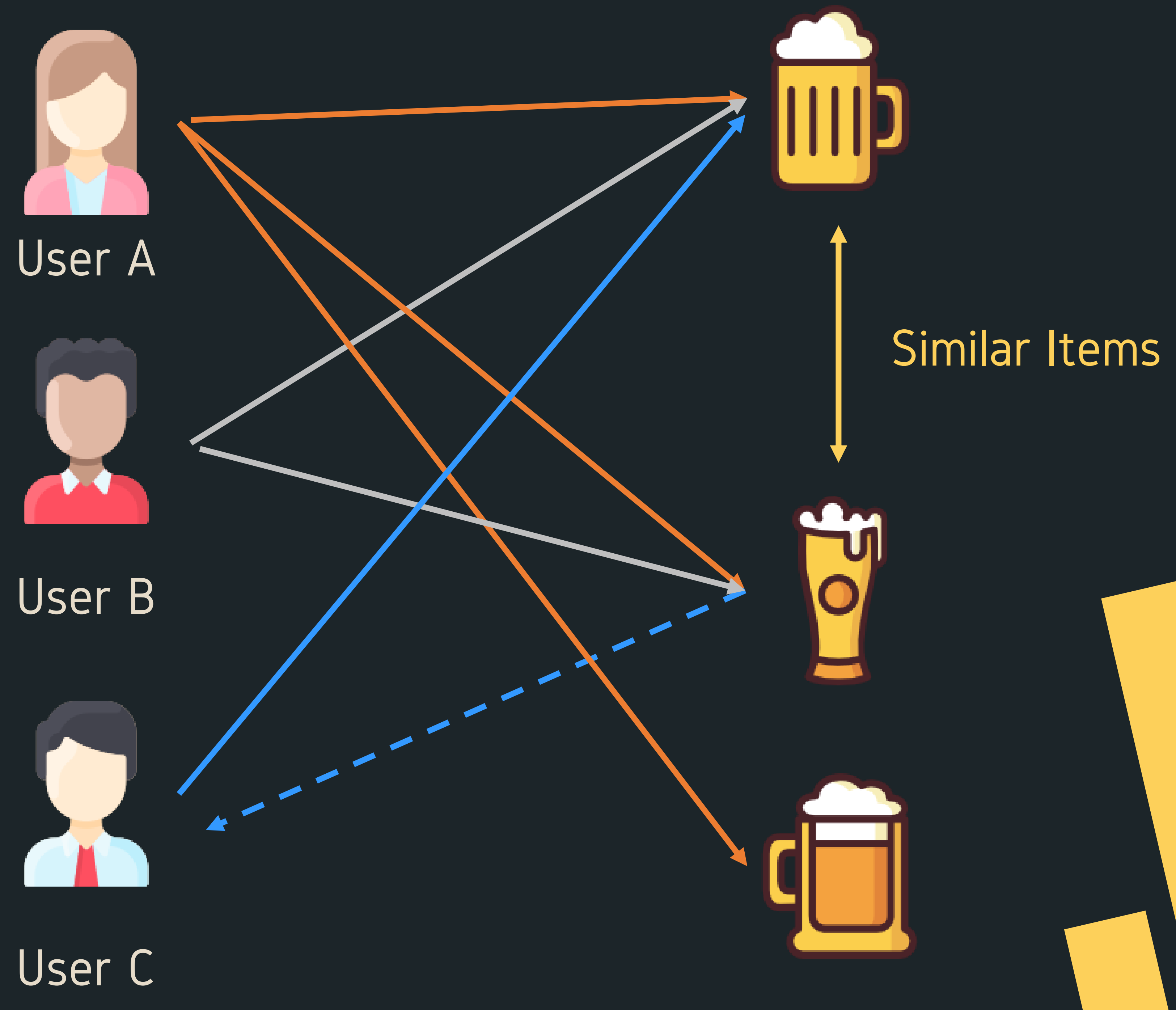
ASSUMPTION

People who like similar items will like items that are liked by other people with similar preferences

USER BASED V.S. ITEM BASED

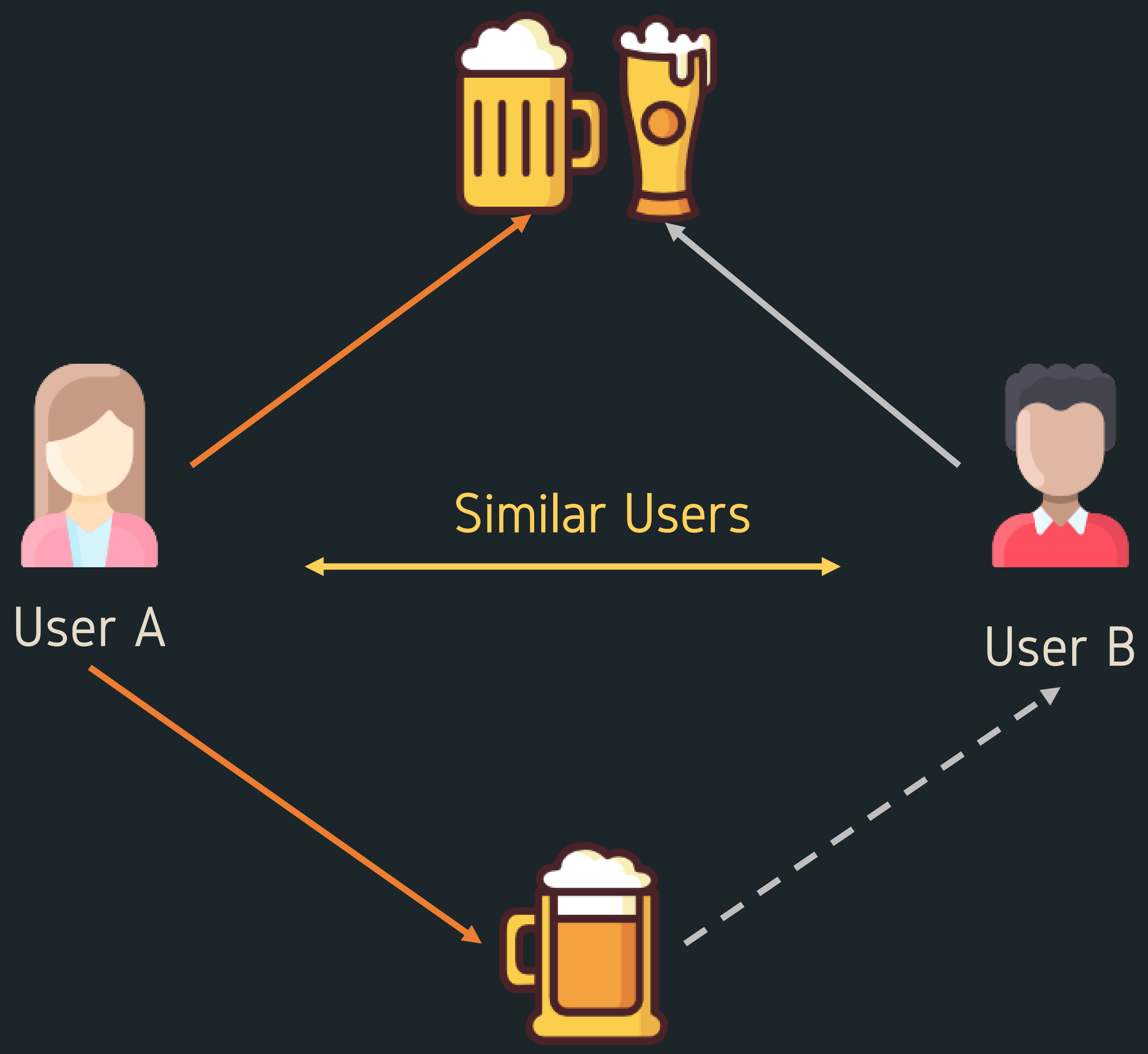


User Based Collaborative Filtering



Item Based Collaborative Filtering

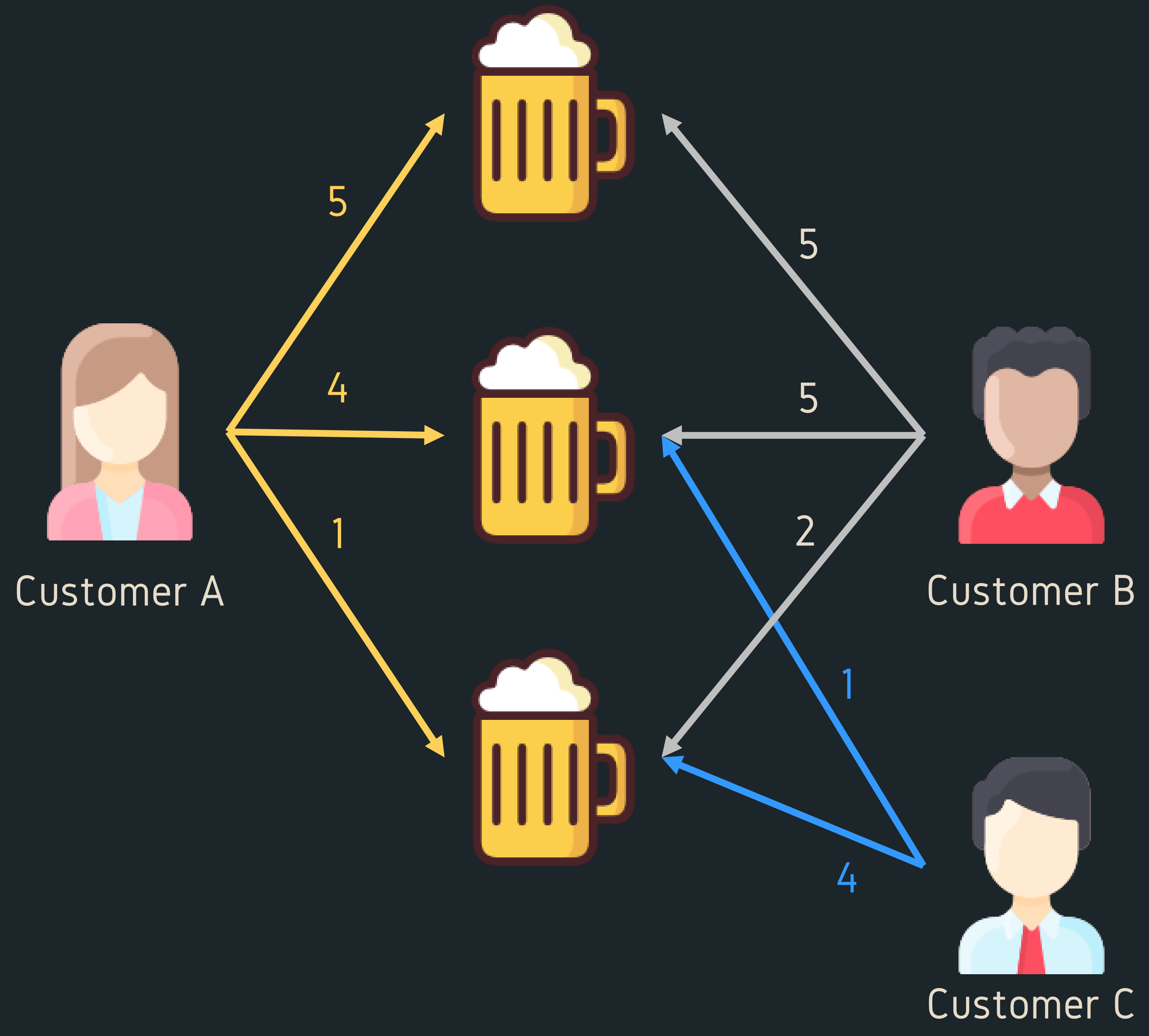
USER BASED V.S. ITEM BASED



User Based Collaborative Filtering

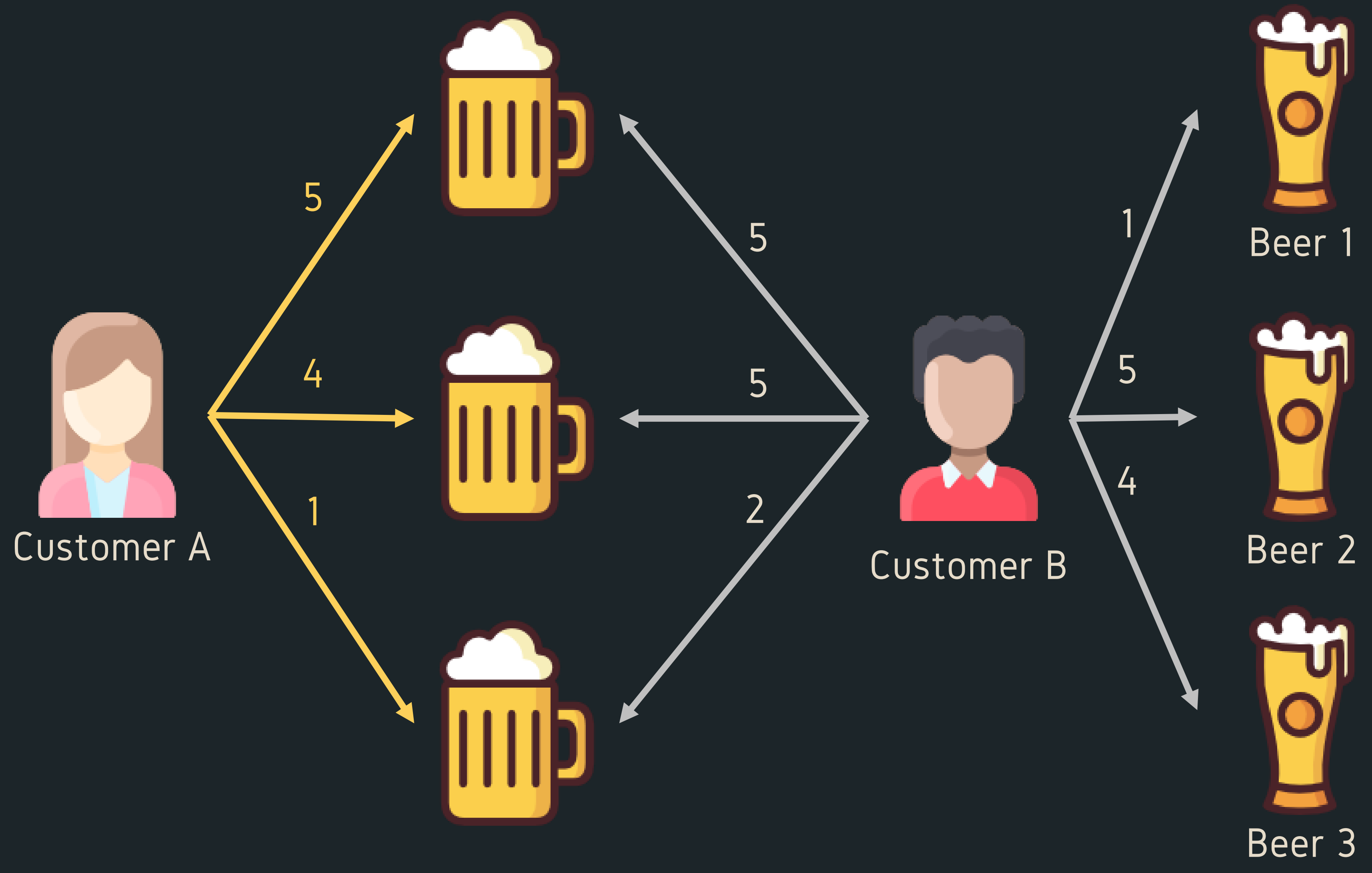
WHY USER BASED?

- Fits our purpose of recommending less commonly found beers
- Gives the users an opportunity to discover new interest in other types of beers



K-Nearest Neighbors (KNN) with Means

1. Calculate similarity between users using **similarity metric**
2. Select the top K similar users



K-Nearest Neighbors (KNN) with Means

3. Take weighted average ratings from these K users, using similarities as weights
4. Rank the items to recommend according to predicted score

PERFORMANCE ANALYSIS

OPTIMAL PARAMETERS WITH GRID SEARCH

- Mean Squared Difference (MSD)
- K=50

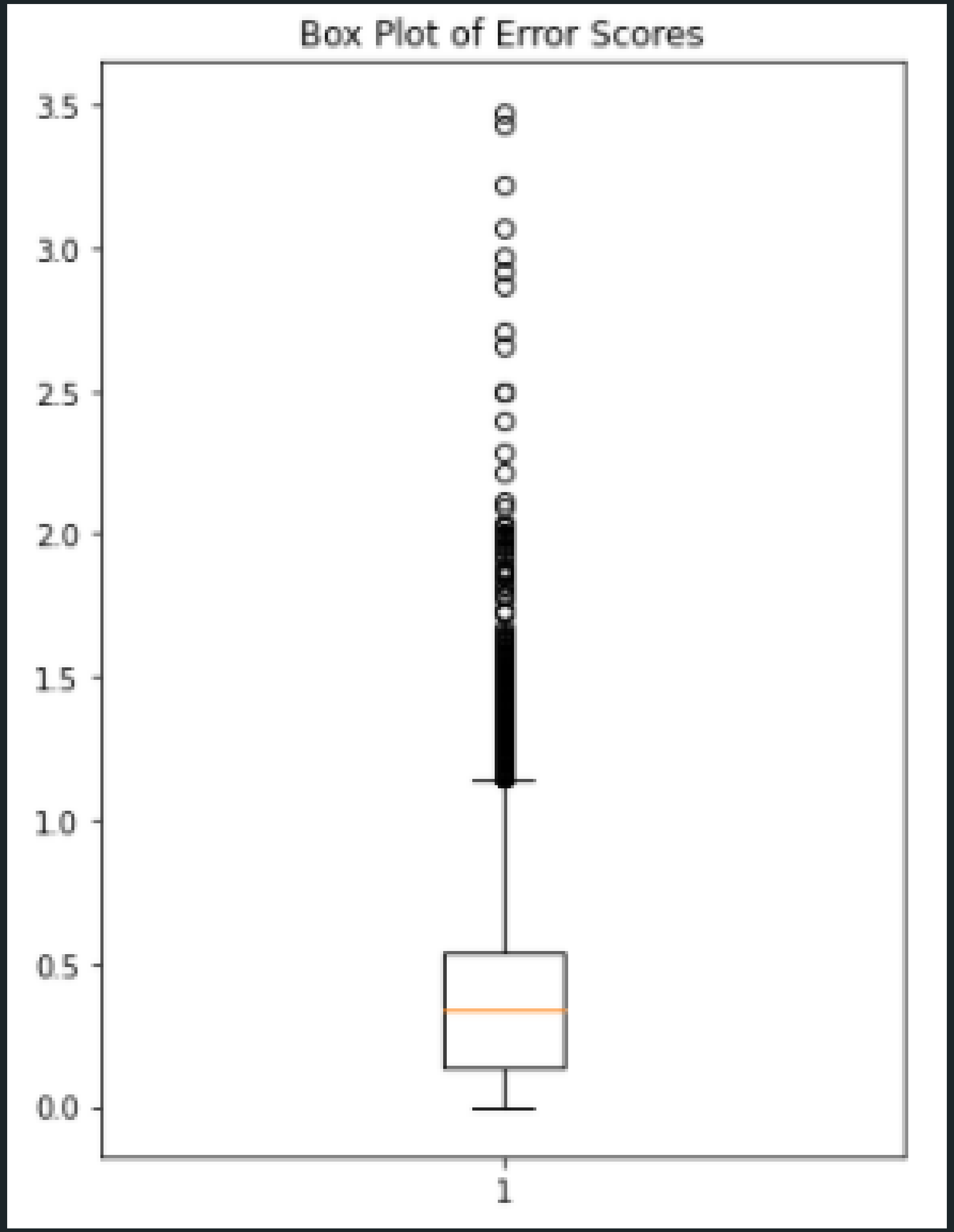
TRAIN TEST SPLIT

- 62,257 Reviews
- Test Set: 30%

PERFORMANCE ANALYSIS

Selects beer with **highest predicted ratings**

Error Metric	Score
Mean Absolute Error (MAE)	0.3992
Root Mean Squared Error (RMSE)	0.2890
Mean Squared Error (MSE)	0.5375



PERFORMANCE ANALYSIS

UNABLE TO PREDICT ACCURATELY FOR NEW USERS

- 1,448 new users out of 6,121 in our test set
- Predicts a score of 4.379839 regardless of beer choice

	user	beer	actual	predict	User Data in Train	error
7	duke1258	Founders Breakfast Stout	5.0	4.379839	0	0.620161
21	sleestak4life	Saison - Brett	4.5	4.379839	0	0.120161
39	Sleestak	Weihenstephaner Hefeweissbier Dunkel	5.0	4.379839	0	0.620161
42	littleg	Tröegs Nugget Nectar	5.0	4.379839	0	0.620161
46	ztprez	Duvel	5.0	4.379839	0	0.620161
...	...	...	...	...	...	...
18666	FlyFisher2782	Temptation	4.5	4.379839	0	0.120161
18669	JayRey	Tröegs Nugget Nectar	5.0	4.379839	0	0.620161
18671	ChuggyMcBeer	Daisy Cutter Pale Ale	5.0	4.379839	0	0.620161
18675	beejayud	Bell's Hopslam Ale	4.5	4.379839	0	0.120161
18677	SpillyBeers	Sculpin India Pale Ale	4.5	4.379839	0	0.120161

LIMITATIONS

Cold Start Problem

Cannot make predictions for users or items that have insufficient information

Dynamic User Preferences

User preferences and item popularity are constantly changing

Similar Taste Ambiguity

Liking popular beer ≠ having similar taste

NEED FOR BETTER RECOMMENDER SYSTEM

NEW USERS

Unable to make recommendation without prior information

CHANGING TRENDS

Unable to dynamically update data

4. MULTI-ARMED BANDIT

WHAT ARE BANDIT ALGORITHMS?

RECOMMENDER SYSTEM

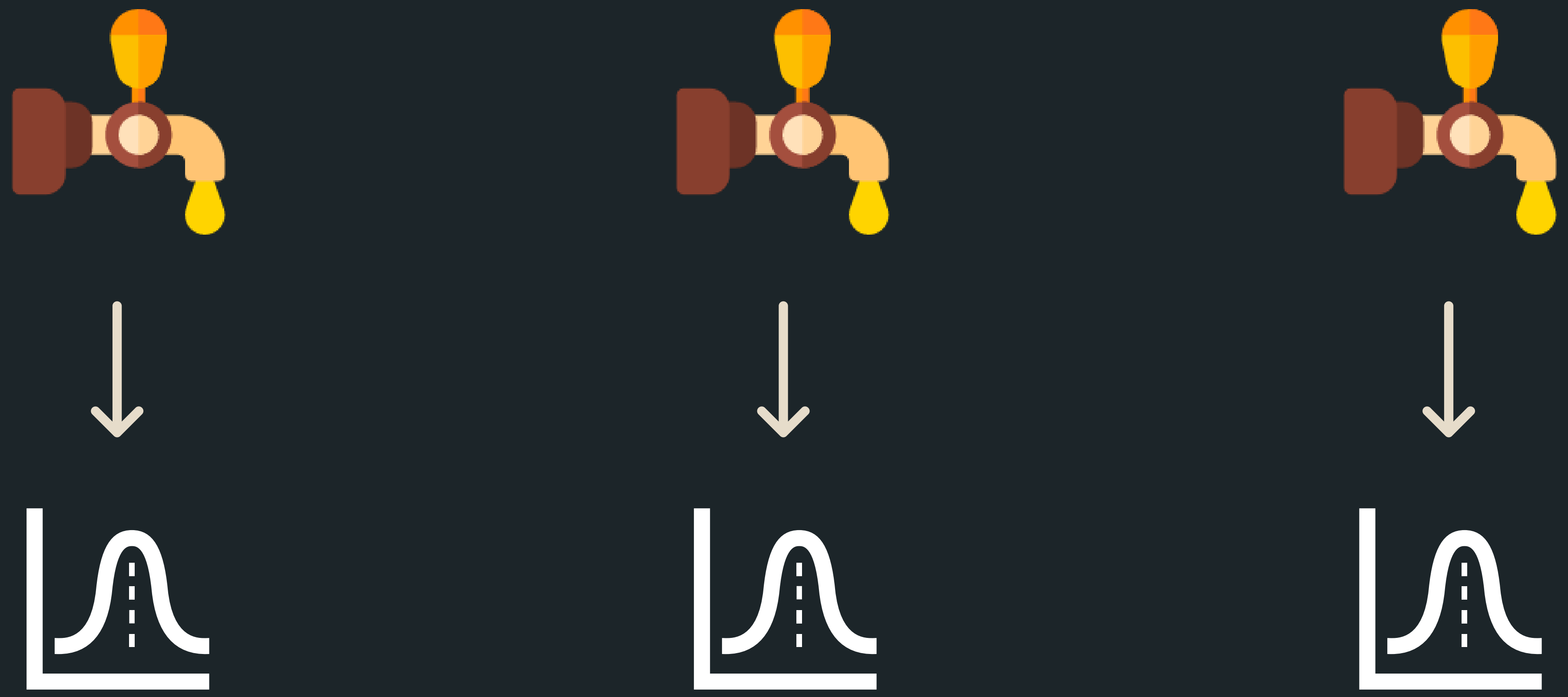
- Maximise rewards
- Trade-off between exploration and exploitation

ANONYMOUS USERS

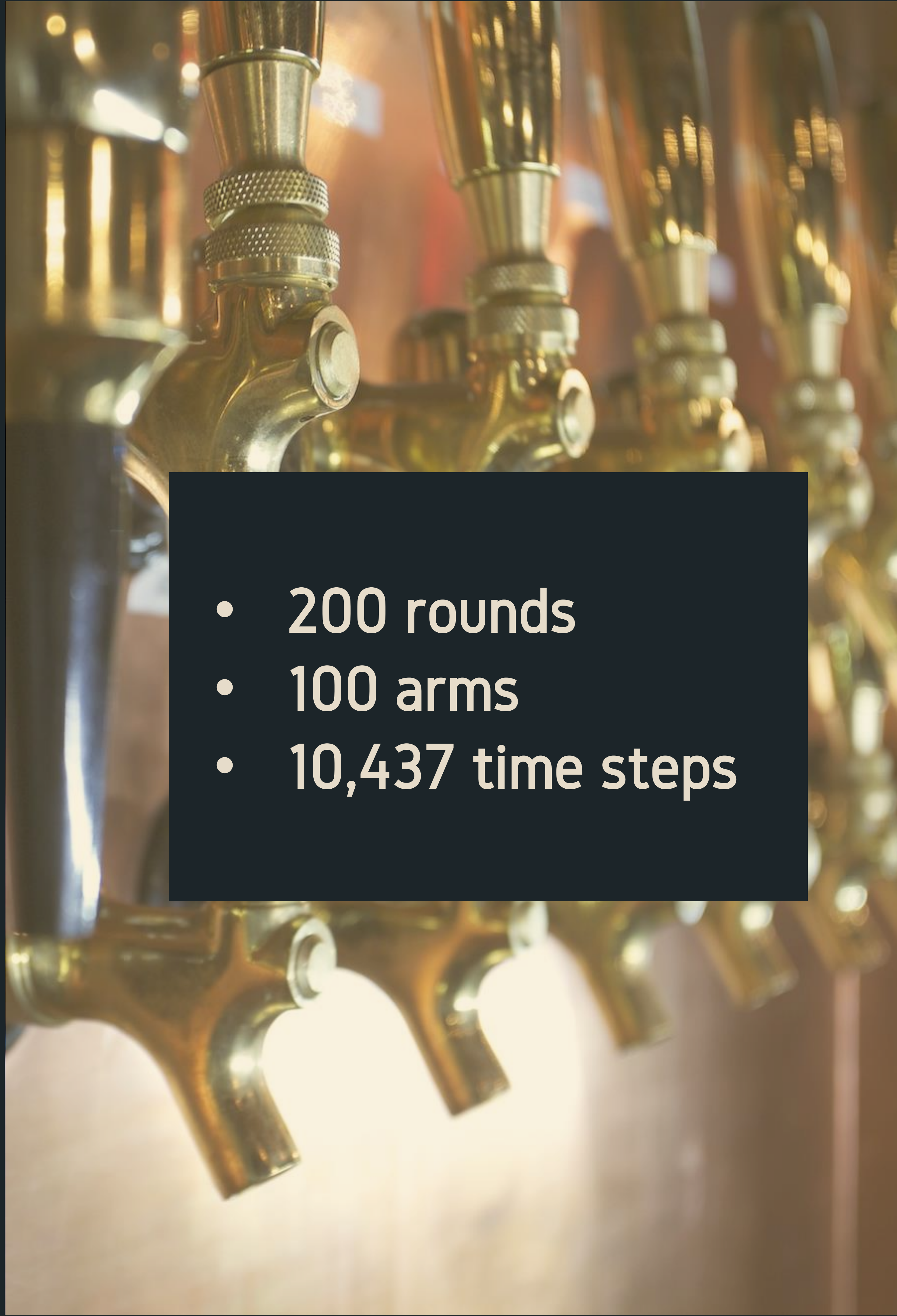
- Do not require knowledge on user or products
- Resolves ‘cold start’ problem

ARM DEFINITION

ARMS



REWARD DISTIBUTION



- 200 rounds
- 100 arms
- 10,437 time steps

BANDIT ALGORITHMS

Epsilon-Decay

Annealing Softmax

UCB1

Bayesian UCB

Thompson Sampling

Epsilon-Decay ε

Epsilon value decays over time

Resources will eventually be diverted to exploit optimal option, rather than to constantly explore the sub-optimal options

Decay function

Hyperparameters:

A : extent of exploration/exploitation

B : slope of transition (middle region)

C : slope at tails

Epsilon-Decay ε

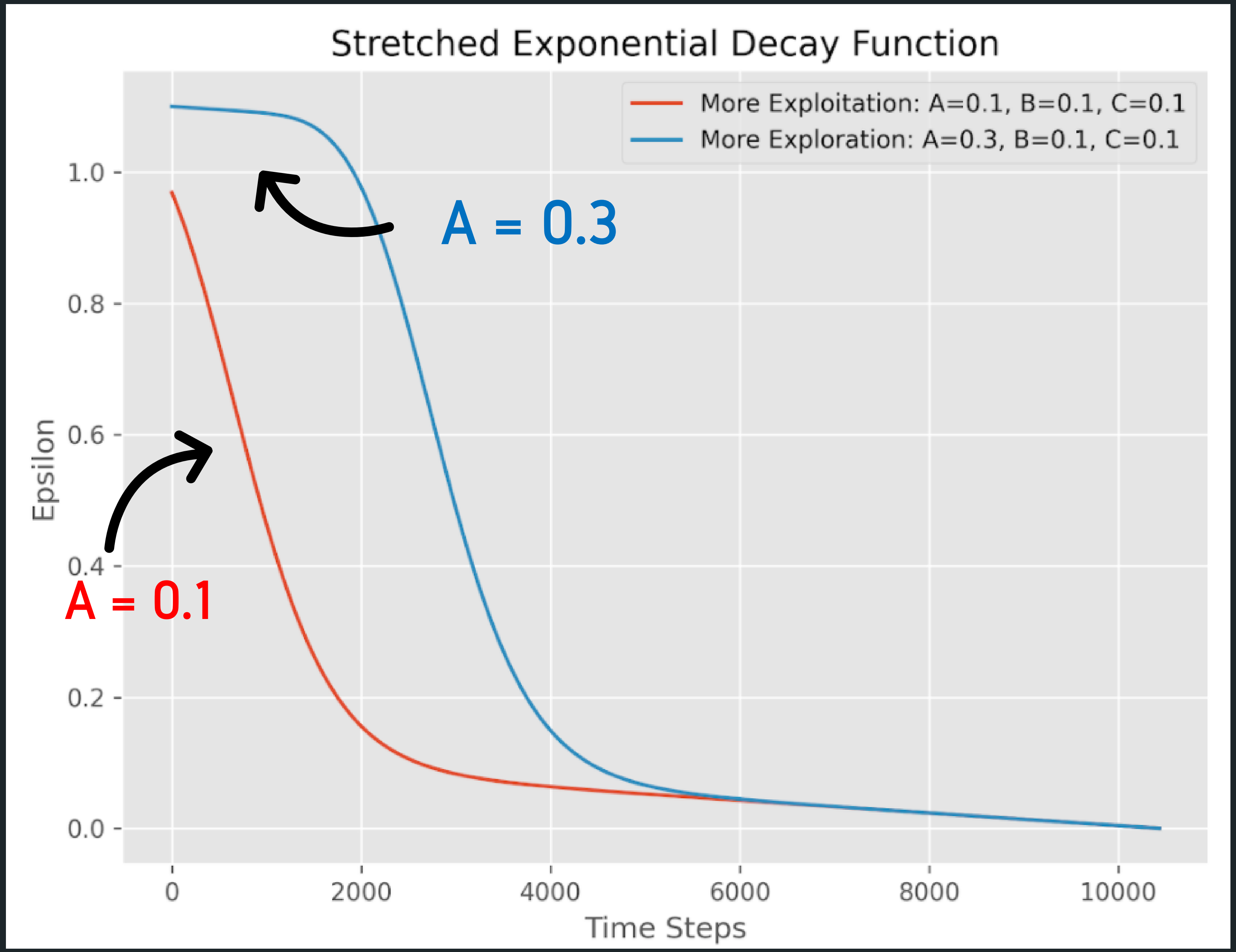
Exploration vs. Exploitation

Exploitation

- A = 0.1
- Epsilon decays faster

Exploration

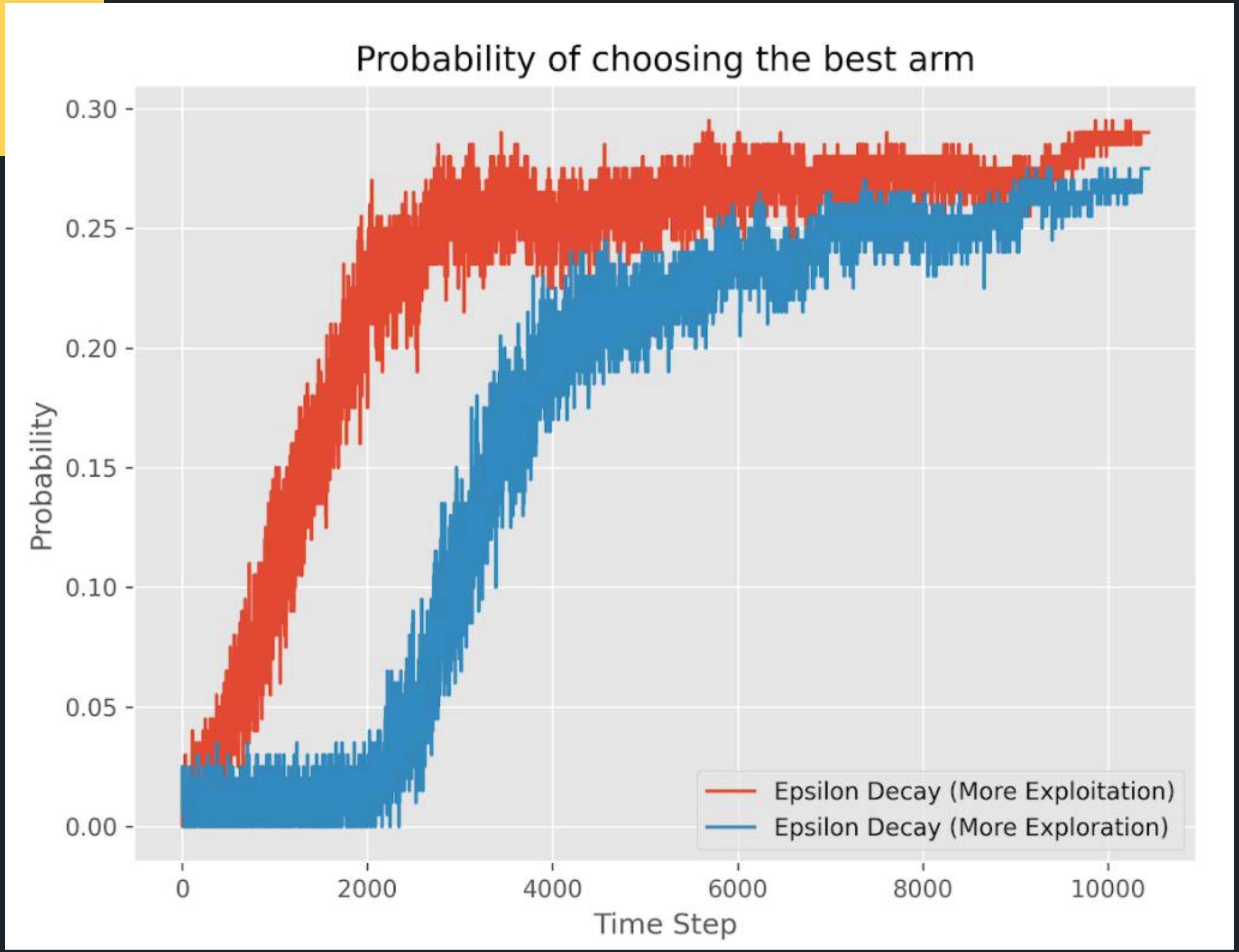
- A = 0.3
- Lengthened left tail
- Epsilon value > 1



PERFORMANCE ANALYSIS

PROBABILITY OF CHOOSING THE BEST ARM

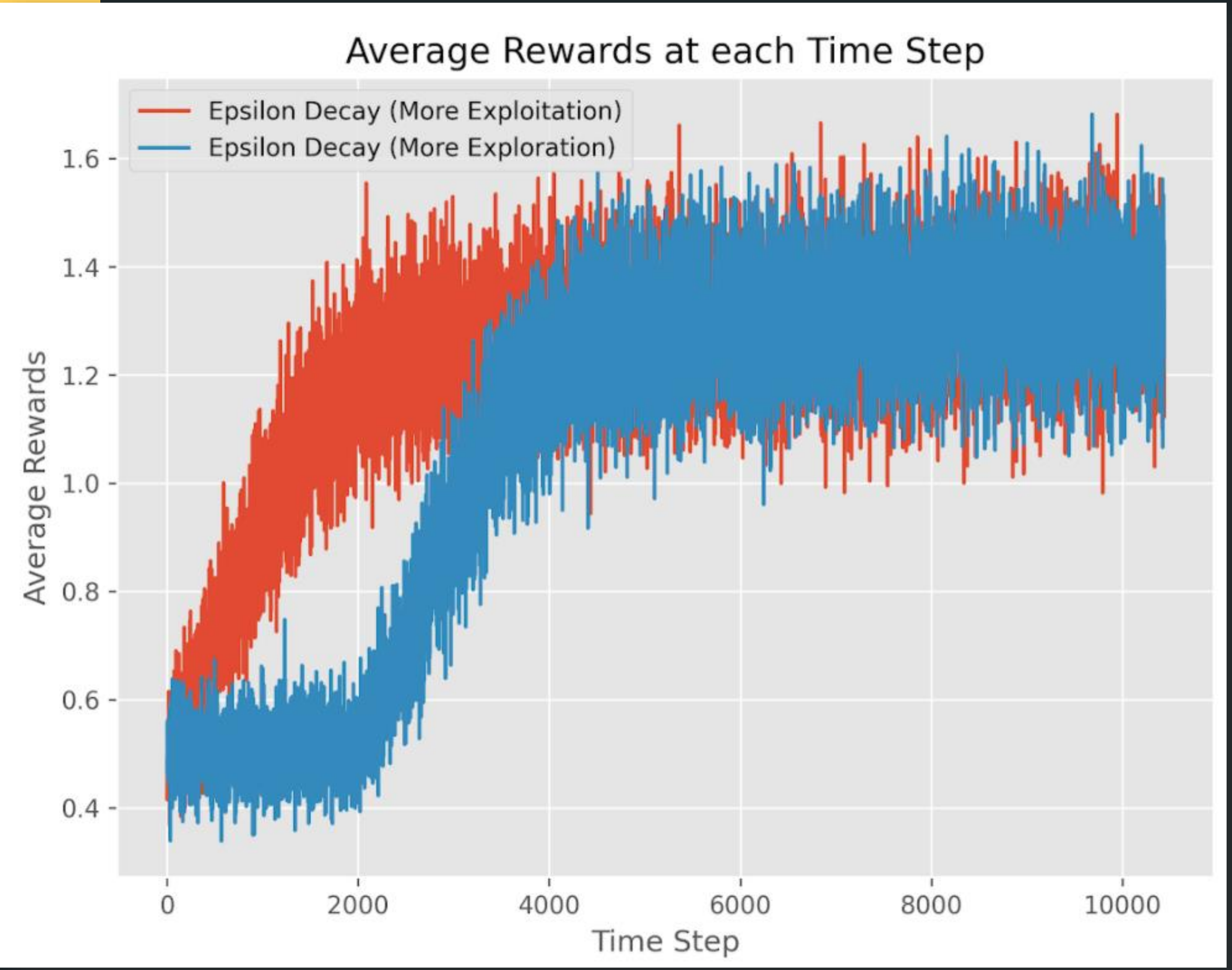
- *Exploitation* converges faster to a higher probability of choosing the best arm



PERFORMANCE ANALYSIS

AVERAGE REWARDS AT EACH TIME STEP

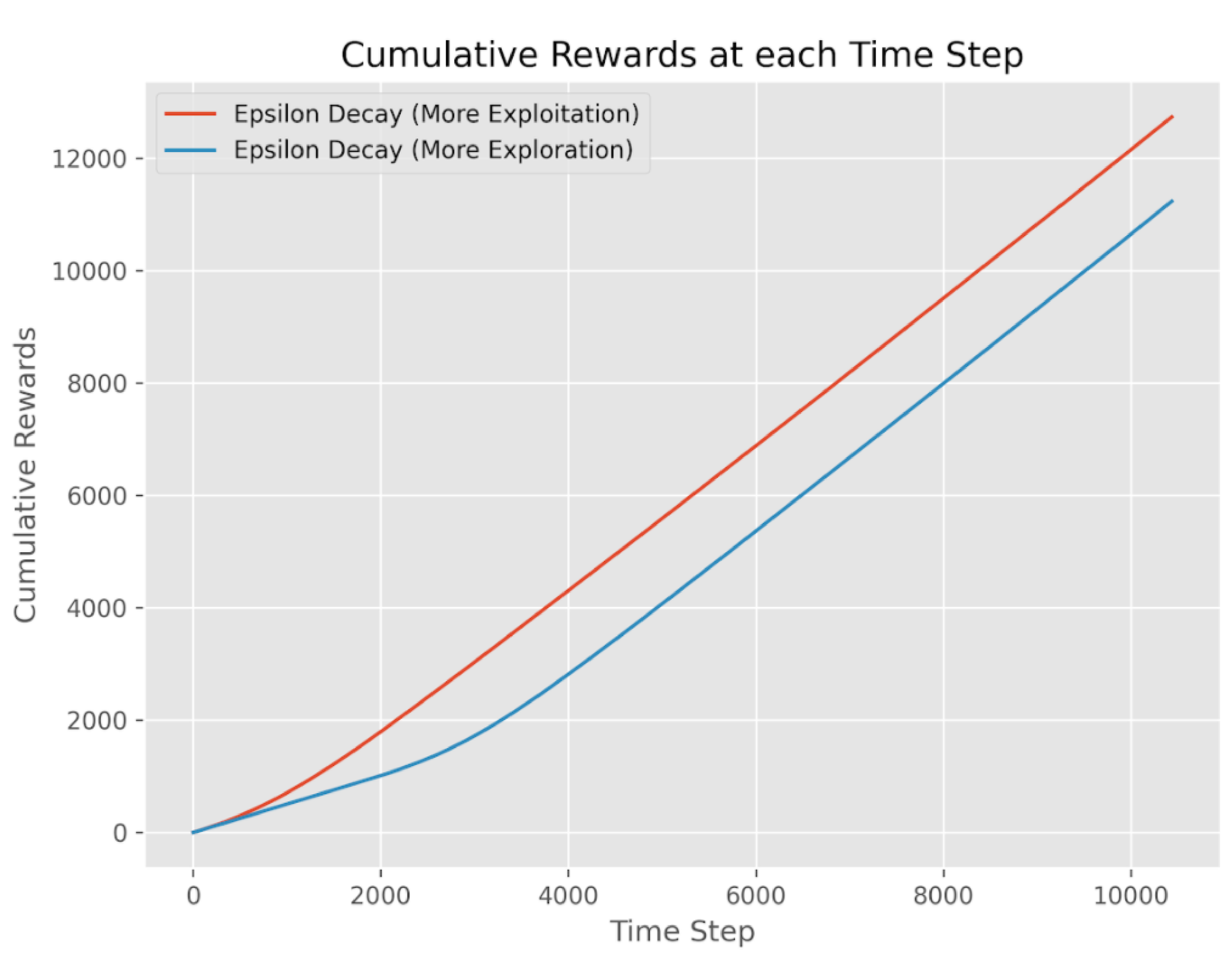
- *Exploitation* converges faster
- Similar short-term performance



PERFORMANCE ANALYSIS

CUMULATIVE REWARDS AT EACH TIME STEP

- *Exploitation* has better long-term performance

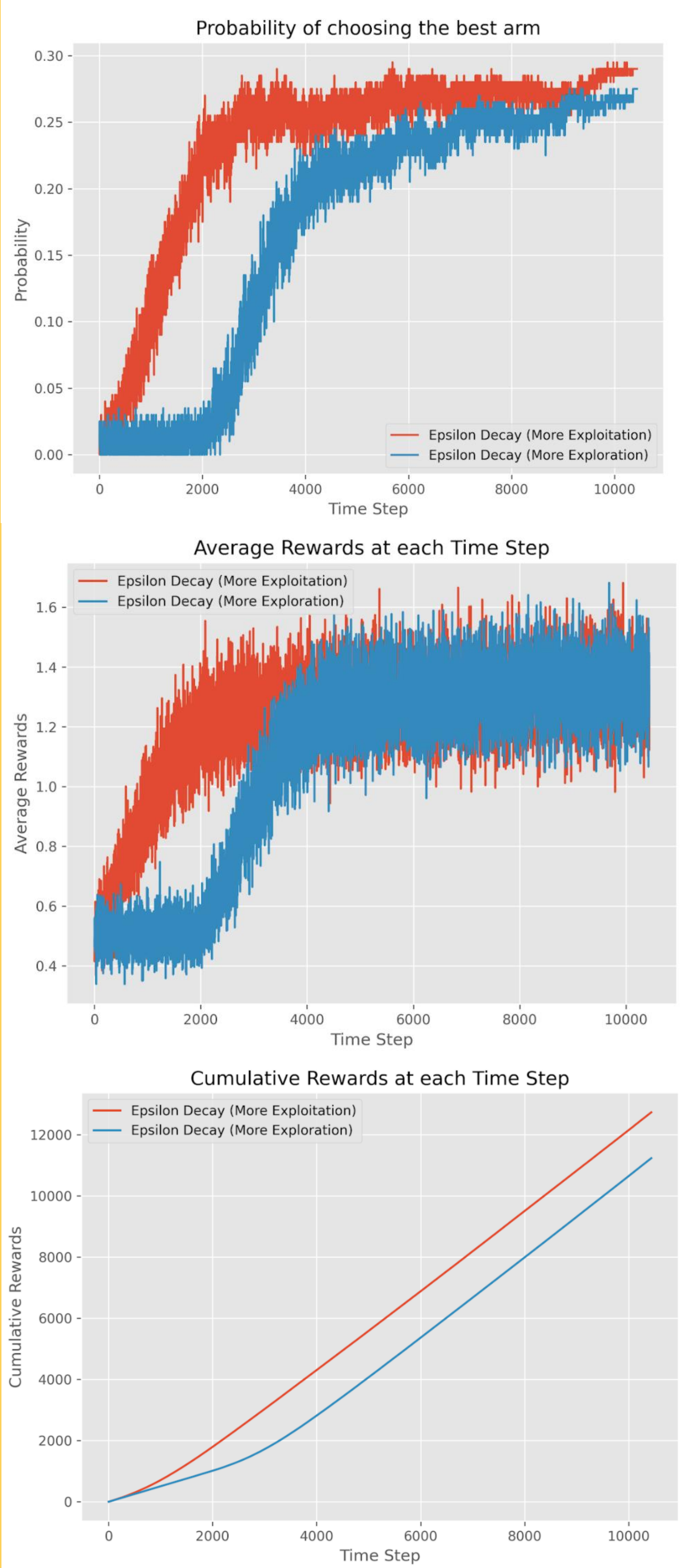


PERFORMANCE ANALYSIS

OVERALL PERFORMANCE

- Comparable short-term performances
- *Exploitation* has better long-term performance

Best Overall Performance:
Exploitation Model



BANDIT ALGORITHMS

Epsilon-Decay

Annealing Softmax

UCB1

Bayesian UCB

Thompson Sampling

Annealing Softmax

Structured Exploration

- Boltzmann distribution
- Arm probability proportional to estimated value
- Temperature parameter *(tau)*

Annealing Factor

- Decreases randomness of selection towards end of simulation
- **Minimise resources wasted**
- **Maximise reward**

BANDIT ALGORITHMS

Epsilon-Decay

Annealing Softmax

UCB1

Bayesian UCB

Thompson Sampling

Upper Confidence Bound (UCB)

- Used both UCB1 and Bayesian UCB algorithms
- Solves the inefficiencies in Epsilon-Decay and Annealing Softmax
- Confidence Intervals introduced considers the knowledge of the arms

Upper Confidence Bound (UCB)

UCB1

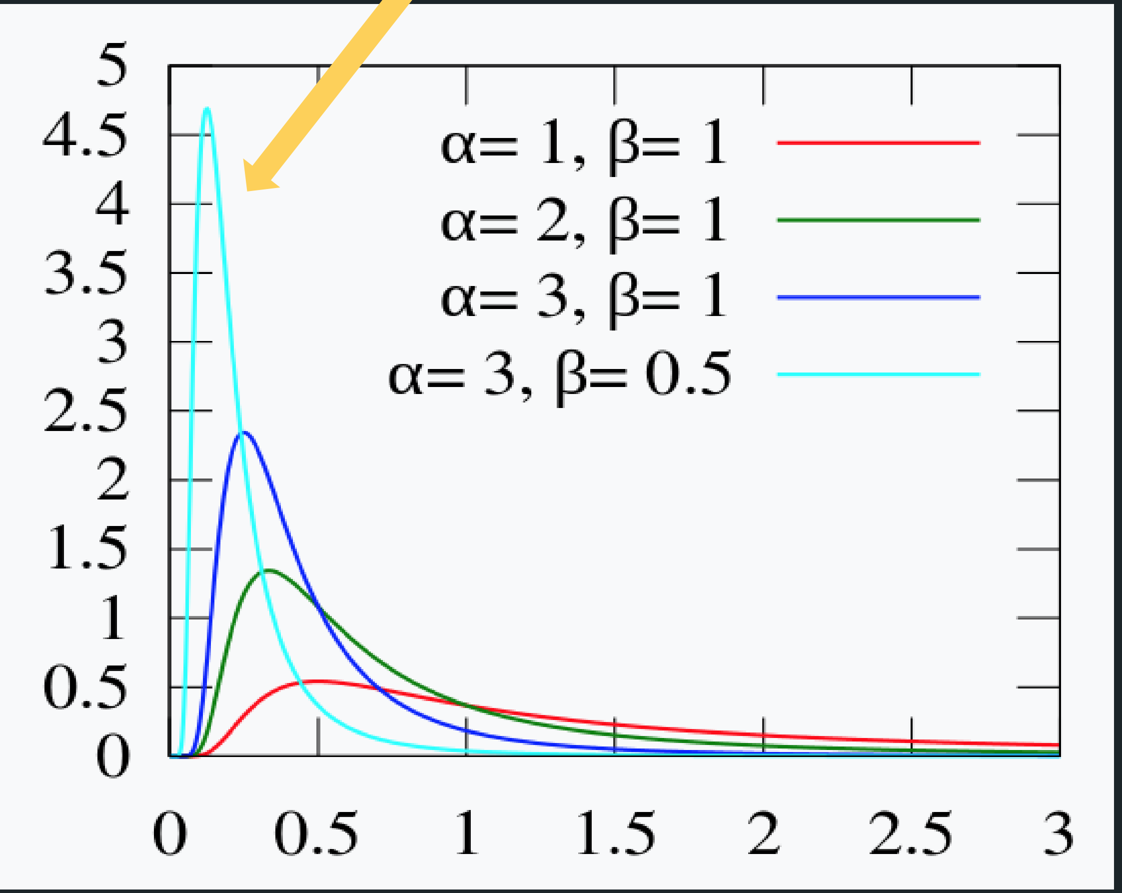
$$A_t = \arg\max_a \left[Q_t(a) + c\sqrt{\frac{\log t}{N_t(a)}} \right]$$

- No prior assumption on the distribution of the arms

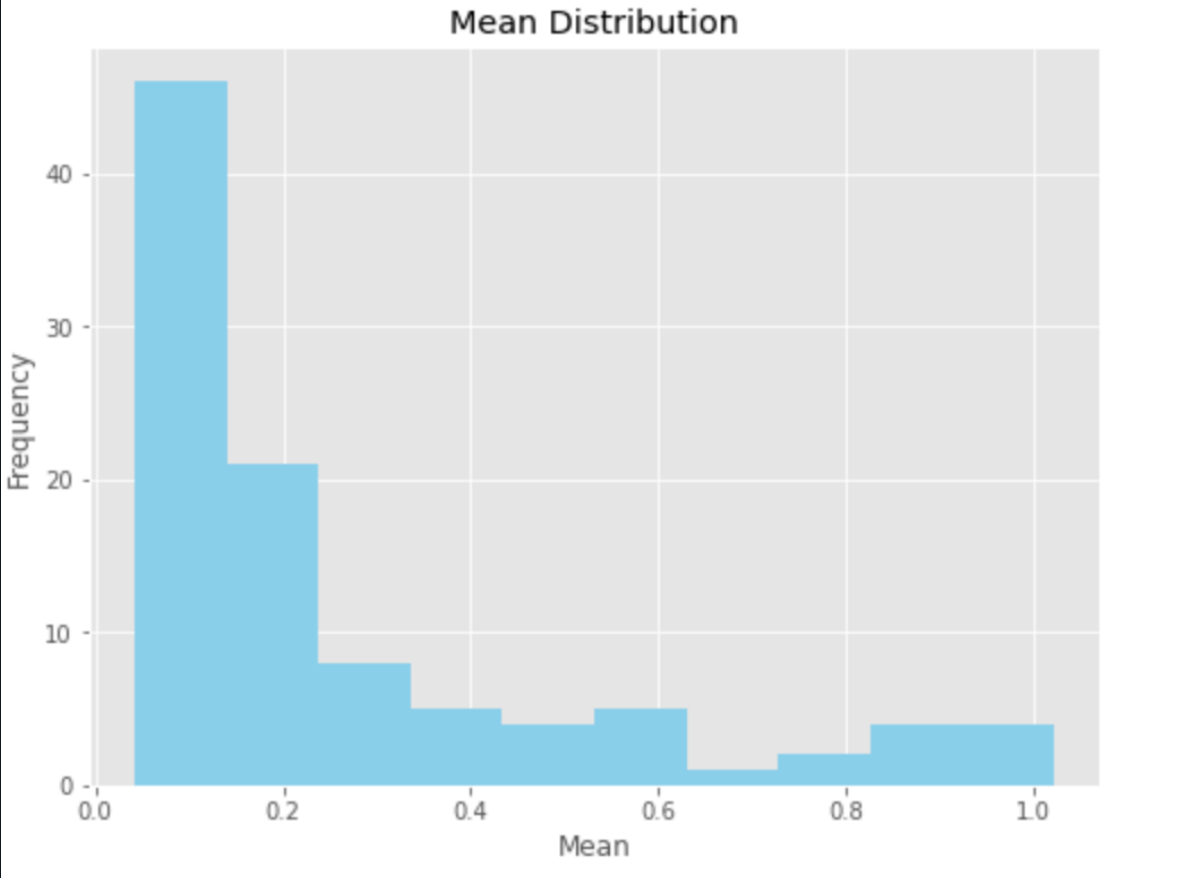
Bayesian UCB

- Incorporates an assumption of the reward distribution of the arms
- Inverse Gamma Distribution used

Upper Confidence Bound (UCB)



Probability Density Function



Distribution of the Mean of Arms

$\alpha = 3, \ \beta = 0.5$

BANDIT ALGORITHMS

Epsilon-Decay

Annealing Softmax

UCB1

Bayesian UCB

Thompson Sampling

Thompson Sampling

Takes Bayesian UCB further

- Samples from a probability model in order to choose the optimal action
- Probability Matching as main strategy

Similarities with Bayesian UCB

- Requires a posterior distribution for the arms
- Utilizes Inverse Gamma Distribution

PERFORMANCE ANALYSIS

Epsilon-Decay

(Exploitation)

Annealing Softmax

UCB1

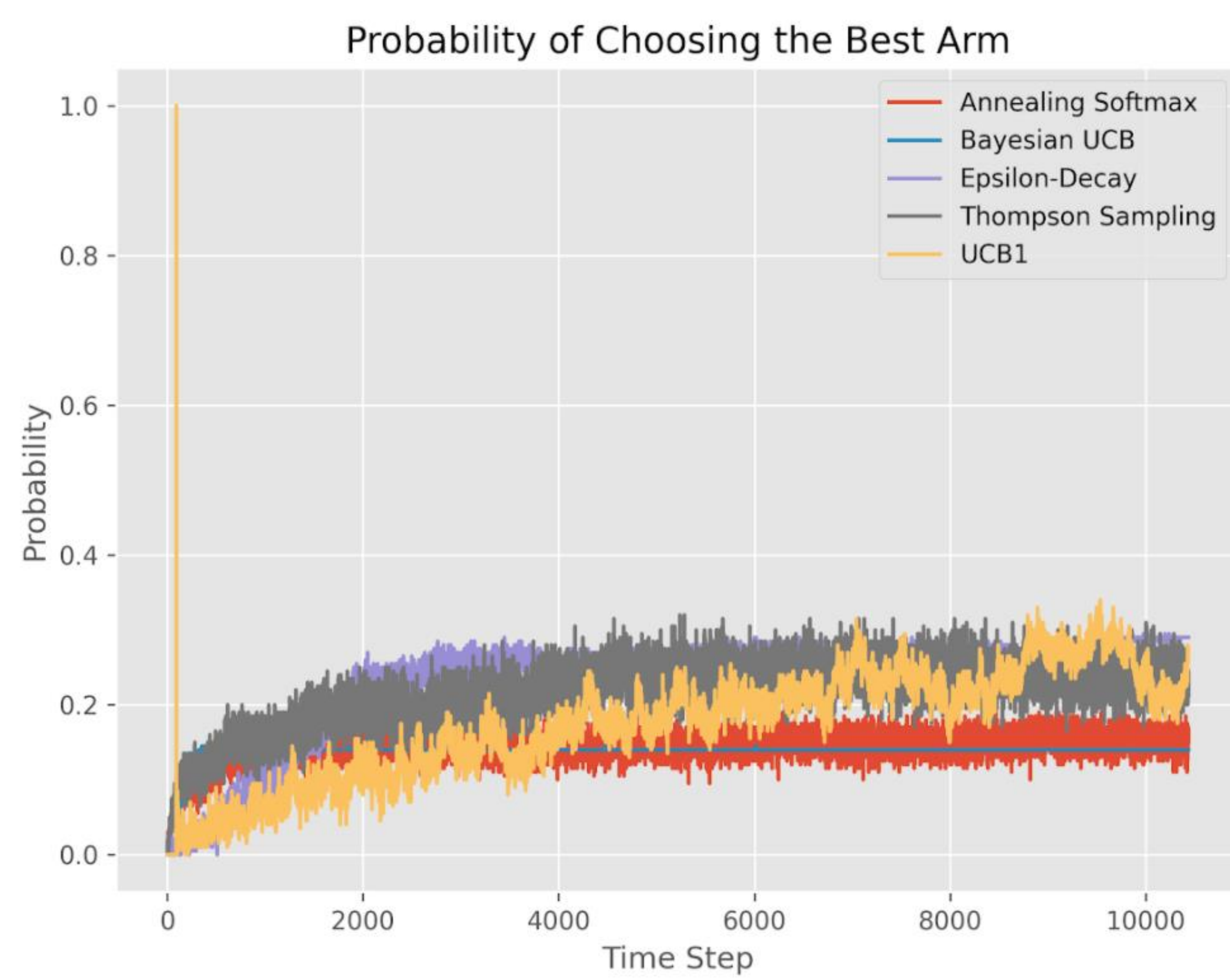
Bayesian UCB

Thompson Sampling

PERFORMANCE ANALYSIS

PROBABILITY OF CHOOSING THE BEST ARM

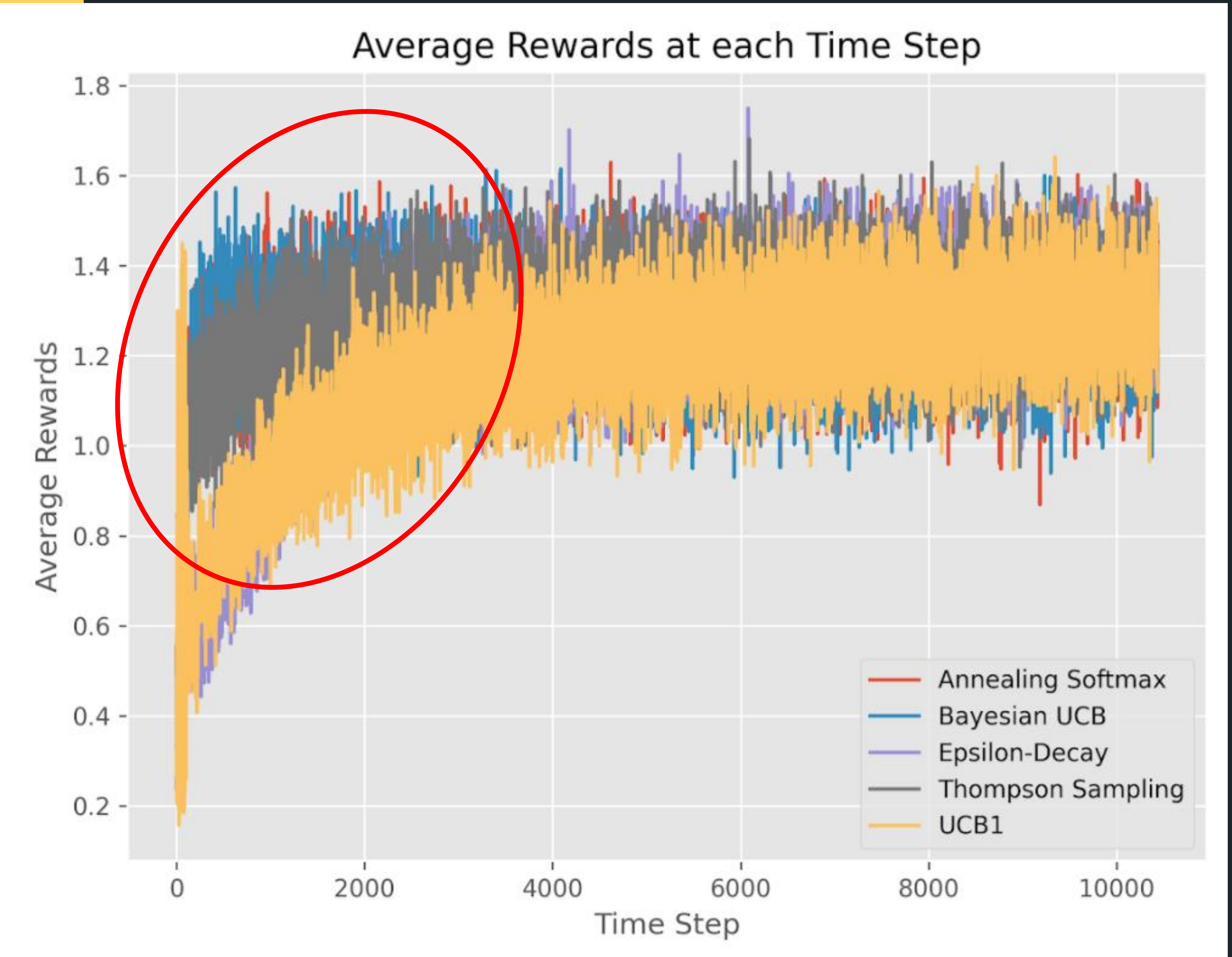
- UCB1, Epsilon-Decay and Thompson Sampling algorithms performs the best
- All converges at same point
- Rate of convergence is highest for Epsilon



PERFORMANCE ANALYSIS

AVERAGE REWARDS AT EACH TIME STEP

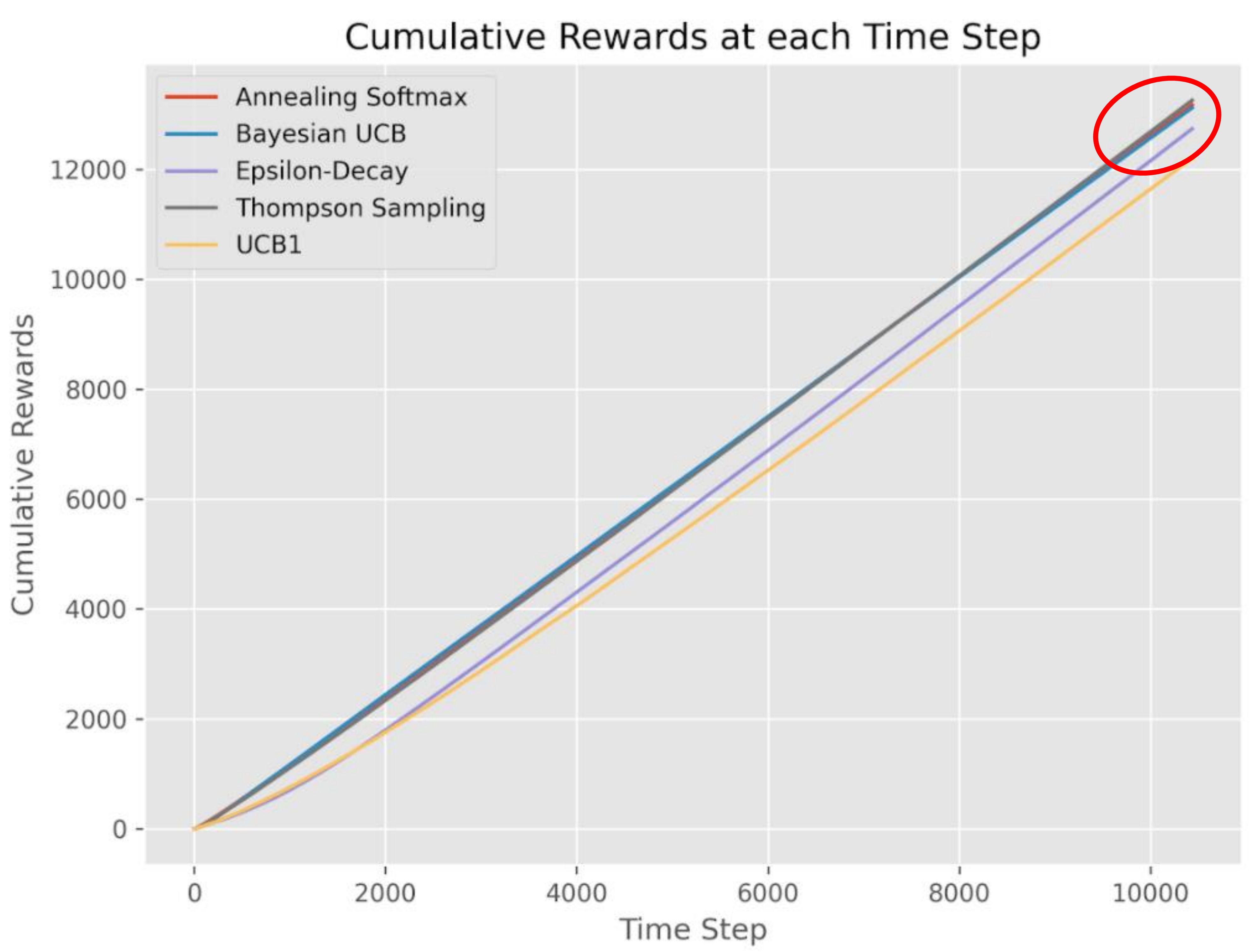
- Peak value is the same for all 5
- Rate of convergence fastest for Annealing Softmax, Bayesian and UCB1 algorithms



PERFORMANCE ANALYSIS

CUMULATIVE REWARDS AT EACH TIME STEP

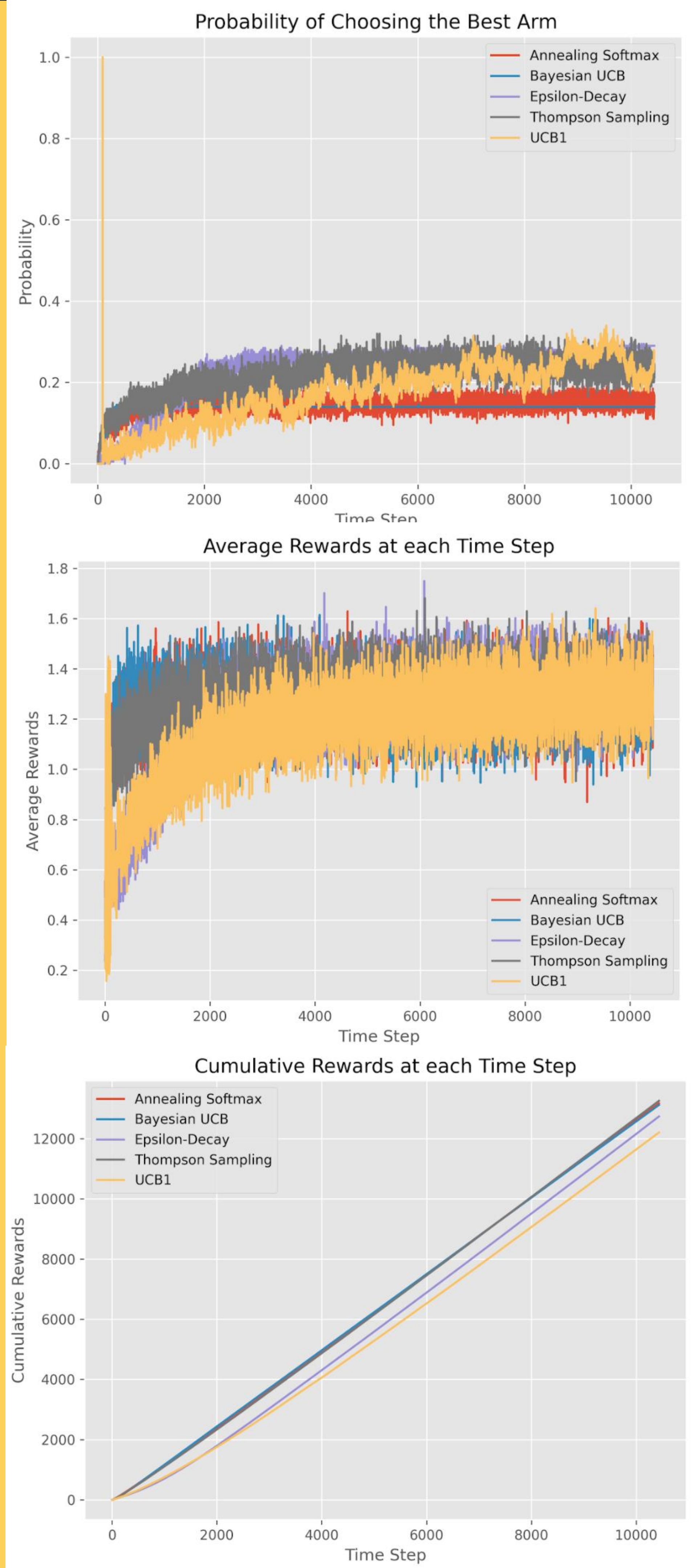
- Thompson Sampling algorithm performed the best
- Annealing Softmax and Bayesian UCB followed extremely closely



PERFORMANCE ANALYSIS

OVERALL PERFORMANCE

- Focus on Cumulative Rewards
- Thompson Sampling, Bayesian UCB and Annealing Softmax have comparable performances



5. ALGORITHM SELECTION

NON- & LOW-ALCOHOLIC BEER

Consumption expected to grow by 31% by 2024

> More health-conscious consumers

> Expect slow fluctuation in beer ratings

ALGORITHM SELECTION

	Epsilon-Decay	Annealing Softmax	UCB1	Bayesian UCB	Thompson Sampling
Use Case Suitability		✓	✓		
Performance		✓		✓	✓

Has **good performance** and **suitable** – can adapt to new trends and maximise profits

Can cause **direct negative impact** to users with **long calibration** required

Takes time to **respond to reward levels changes** and **select the new best-performing option**

6. FUTURE IMPROVEMENTS

FUTURE IMPROVEMENTS

OBTAINING AN UPDATED DATASET

- Helps to enhance the system accuracy

TEST OUT DIFFERENT METRICS/ALGORITHMS

- Using conversion rate or purchase volume
- Applying adversarial or contextual bandit
- More accurate real-world simulation and recommendations

7. CONCLUSION

CONCLUSION

ERA OF BIG DATA ANALYTICS

Discover insights, improve decision-making processes and increase efficiency

BUILDING A BEER RECOMMENDER SYSTEM

Using the Annealing Softmax algorithm to recommend less commonly found but high-quality beers

VALUE DELIVERED

Minimise inventory loss, increase consumer engagement and differentiate our business from competitors

THANK YOU!

Keep in touch

TAN ZEN WEI
zenwei.tan@u.nus.edu

CHUA KAI BING
kai_bing.chua@u.nus.edu

CHAN CHEAH CHA
chan.cheahcha@u.nus.edu

LIM JIA QI
e0325210@u.nus.edu

GOH JIA YI, JESA
gohjiayi@u.nus.edu