

Description

Multi armed bandit (MAB) is a popular problem in reinforcement learning paradigm. Multi-armed bandit problems are the most basic examples of sequential decision problems with an exploration-exploitation trade-off. This is the balance between staying with the option that gave highest rewards in the past and exploring new options that might give higher rewards in the future. There are many different variants of MAB problems which can in terms of reward structure (stochastic, adversarial and markov rewards), feedback structure (bandit and semi-bandit), information available (contexts or context free), availability of arms (sleeping MAB) etc. MAB finds it application in various domains like Sponsored search auctions, crowdsourcing, demand response in smart grids, routing etc. It is famously used by Google Analytics e.g. [1,2].

Typically, the performance of MAB algorithms is evaluated in terms of regret (definition may vary with different variants of MAB). There are many algorithms in literature which addresses the MAB problem. Some of the well known algorithms are Upper Confidence Bound (UCB) and Thompson Sampling.

Goal

The aim of the project is to build a tool with an attractive UI where a user can upload the data and select the algorithm and tool outputs regret which may be in a file or as a graph. The motive of this project is to provide abstraction in the running the algorithm where the user doesn't need to get into details of algorithms implementation.

The user should be able to

- 1) Pick required algorithms for comparison/evaluation
- 2) Upload/Enter/generate required input data
- 3) Get all graphs/output for comparison/evaluation

Tasks to be done

- 1) A python kind of package with a decent UI
- 2) Implement given set of MAB algorithms (epsilon greedy, Boltzmann Exploration (Softmax), Pursuit Algorithms, Reinforcement Comparison, UCB1, EXP3, Thompson Sampling, KL UCB, LinUCB) and fine tune them.
- 3) Plot regret and compare different algorithms for the given problems.
- 4) Stretch Goal: Come up with new MAB algorithm for newer settings

Expectations

- 1) Understand theory behind MAB algorithms (e.g. [3]).
- 2) Python and UI development

3) Understanding state of the art MAB algorithms [4]

References

1. Akash Das Sarma, Sujit Gujar, Y Narahari, "Truthful Multi-Armed Bandit Mechanisms for Multi-Slot Sponsored Search Auctions". *Current Science, Special Issue on Game Theory*, Volume 103(9), pp 1064-1077, November 2012.
2. Shweta Jain, Sujit Gujar, Satyanath Bhat, Onno Zoeter, Y. Narahari, "A Quality Assuring, Cost Optimal Multi-Armed Bandit Mechanism for Expertsourcing". *Journal of Artificial Intelligence (AI)* 254 (2018): 44-63.
3. Abhishek Kumar, Sneha Maheshwari, and Sujit Gujar. "Introduction to Concentration Inequalities." *arXiv preprint arXiv:1910.02884* (2019).
4. Volodymyr Kuleshov, Doina Precup "Algorithms for the multi-armed bandit problem" *Journal of Machine Learning Research* 1 (2000) 1-48.