

Forschungspraxis

# Multi-armed Bandits - Basic algorithms and possible applications

Multi-armed bandits is a simple but very powerful framework for algorithms that make decisions over time under uncertainty.

In the basic version, an algorithm has  $K$  possible actions to choose from, a.k.a. arms, and  $T$  rounds. In each round, the algorithm chooses an arm

and collects a reward for this arm. The reward is drawn independently from some distribution which is fixed (i.e., depends only on the chosen arm), but not known to the algorithm.

we have a tradeoff between exploration and exploitation: making optimal near-term decisions based on the available information. This tradeoff, which arises in numerous application scenarios, is essential in multi-armed bandits. Essentially, the algorithm strives to learn which arms are best (perhaps approximately so), while not spending too much time exploring.

The plan is to study the basic algorithms and implement them using relevant real-life data.

References: The following two books are comprehensive surveys of the field. Fortunately, both are freely available online:

(1) Alexander Slivkins, Introduction to Multi-armed Bandits, 2019.

<https://arxiv.org/abs/1904.07272>

(2) T. Lattimore and C. Szepesvari, Bandit Algorithms (Cambridge University Press, 2020)

<https://tor-lattimore.com/downloads/book/book.pdf>

## Prerequisites

Interest and some background in Probability, Information theory and Algorithm design.

References

## Advisors

Abdalla Ibrahim