K-Armed Bandit based approach to Recommendation Systems

- Srivatsan Srinivasan (21011101129)

Problem Formulation:

In the context of a recommendation system, each arm (‘k’) represents a different recommendation (e.g., items, movies, products, etc.). The goal is to maximize user engagement (e.g., clicks, purchases) by selecting the best recommendation.

Defining Variables

Arms: Each possible recommendation option.
Rewards: The feedback received from the user (e.g., 1 for a click, 0 for no click).
Action Selection: The process of choosing which recommendation to show to the user.
Exploration: Trying out different recommendations to discover their potential.
Exploitation: Leveraging the best-known recommendation to maximize rewards.

Problem Solution Overview

Teaching the RL Model can be based on the *“Epsilon Greedy Algorithm”*

Where in,

With a probability epsilon, we can choose a Random “Arm” (Recommendation) for the purpose of Exploration.
Thereby, with a probability of 1 - epsilon, we can choose an Arm that is highly rewarded thus far (optimal thus far) to serve the purpose of Exploitation.
The value of “epsilon” decays over time.