

K-Arm Bandit Solution for News Value Maximizer

Introduction:

In the digital media landscape, politically and commercially affiliated media companies often aim to maximize the visibility of certain articles that align with their objectives or values. A K-arm bandit approach provides an effective way to dynamically prioritize these articles based on their views and engagement, balancing exploration (testing new articles) and exploitation (focusing on popular, high-value articles).

Problem Definition:

The task is to implement a K-arm bandit algorithm that can continuously choose articles to maximize views for a set of aligned articles. Each 'arm' represents a different article, and the algorithm aims to maximize the cumulative reward (views) for aligned articles over time. This model adjusts based on article performance, allowing it to dynamically adapt to changes in reader interest.

Approach:

This solution uses an **epsilon-greedy K-arm bandit approach** where each article has an associated reward function based on past views and engagement rates. The strategy begins with a high exploration rate to learn the initial popularity of each article and gradually transitions to exploitation, focusing on articles with the highest average views.

1. Initialization: Each article is assigned a reward counter and an engagement score.
2. Epsilon-Greedy Strategy: At each time step, the algorithm:
 - Selects a random article with probability epsilon (exploration).
 - Selects the highest-performing article with probability $1 - \text{epsilon}$ (exploitation).
3. Reward Update: After each selection, the views for the chosen article are recorded as a reward, and the reward counters are updated.
4. Dynamic Adjustment: Epsilon decreases over time to favor exploitation as high-value articles are identified.

Algorithm Implementation (Pseudocode):-

1. Initialize:

- Set epsilon to a starting value (e.g., 0.1).
- Assign each article an initial score of 0.

2. For each article selection:

- With probability epsilon, select a random article.
- Otherwise, select the article with the highest average views.
- Record the views of the chosen article as the reward.

3. Update:

- Add the new reward to the article's total views and update its average views.
- Gradually decrease epsilon to favor exploitation over time.

Expected Outcome:

The algorithm will initially explore a range of articles, including newly published or under-viewed ones, to gauge their engagement potential. Over time, it will increasingly focus on promoting high-performing articles that align with the company's objectives, maximizing total views and ensuring that prioritized content receives more visibility.

This approach ensures that both newly introduced and consistently popular articles are given opportunities for visibility while optimizing for the aligned content views.