

## REINFORCEMENT LEARNING

### K-arm bandit based solution for A news value maximiser:

**Politically and commercially affiliated media companies are tasked with maximizing the views for certain articles more than others. Build a system that maximises the views for these “aligned” articles.**

- Media companies need to maximize views on certain politically or commercially aligned articles.

- Treat each article as an “arm” in a K-armed bandit problem.
- Reward = number of views the article gets after promotion.
- Need to balance between exploring new articles (randomly promoting) and exploiting the best-known articles.
- Use Epsilon-Greedy:
  - With probability  $\epsilon$ , pick a random article.
  - With probability  $1-\epsilon$ , pick the article with the highest reward so far.
  - After promoting an article, update the strategy based on the views it got.
- Over time, this approach helps focus on articles that consistently get more views, maximizing the total view count for aligned articles.

Nakshatiraa K N

21011101080