Movie Recommendation System using K-Arm Bandit Problem

-Shreya Sevelar
21011101123

- Each possible movie is modelled as an arm

- Out of 100 movies (k = 100) the system presents 10 choices to the user. The user selects a movie based on which the expected reward for that movie is updated.

- We use the epsilon greedy approach to control exploration and exploitation.

Initialisation:

- We have an array, N which is set to all zeroes – stores the number of times a movie has been chosen.

- We have an array, Q set to all zeroes – this is the expected reward for each movie

- Let epsilon = 0.3

Training:

- Generate a random number r between 0 and 1.

- If r<0.3:

  - Exploration: select 10 movies at random out of the 100 and present them as choices to the user

- If r>=0.3:

  - Exploitation: Select 10 movies based on the expected reward array Q. We do this by choosing the movies with the highest values of Q; i.e top 10 values from Q.

- **User Choice:**

  - User chooses one of the 10 movies

  - The number of times that movie has been chosen is incremented by 1 in the N array : Nmovie+=1

  - We now update the expected reward for that movie, Q using the following formula:
    Qnew = Qold + (R-Qold)/Nmovie

This process is repeated for T timesteps.