RL CIA-1

Problem statement: Develop a movie recommendation system using the K-arm bandit algorithm to enhance user viewership  
  
 Here each arm represents a different movie that could be recommended to the user. The reward is based on whether the user clicks on the recommended movie and watches it for a certain amount of time ( Say if he/she watches the movie for atleast 10 minutes, the reward is 1 else 0). The recommendation system uses Epsilon-Greedy algorithm to choose which movie to recommend. As system does't know what genre of movies user like, it explores different movies. Eventually using obtained rewars, the system starts to recommend movies that user like. The system adapts to the specific user's preferences over time. By balancing exploration and exploitation, the system not only recommends what the user has already shown interest in but also tries new options that might be even better.