**Multi-Armed Recommender System Bandit Ensembles**

**Bandit Recommender System Ensembles**

Recommendation is a cyclic interactive process

- **External sources:**
  - Search
  - Browsing
  - Questionnaires
  - Etc.
- **Input data (observations)**
- **Learning (exploration)**
- **Bandit Recommender System Ensembles**

**Experimental Approach**

1. Initial random rating split
   - Each system has its own training and test sets (only the initial sets are common to all systems)
2. The system produces a recommendation for every user
3. After each epoch:
   - Compute-cumulative recall with test set
   - Add users’ feedback to training set and go back to step 2

**Three types of systems: bandit ensembles, standalone algorithms and a dynamic non-bandit ensemble**

1. Select an arm according to the estimated reward
2. Obtain reward returned by the selected arm
3. Update estimated reward model of selected arm

**Results**

- **Cumulative recall curves**
- **Traffic ratio of bandit ensembles**
- **Feedback loop in the dynamic non-bandit ensemble**

**References**