# User generated Recommendation System using Knowledge based System

Prashant Das
SCOE, Mumbai University,
Mumbai, India

Apurva Bansode
SCOE, Mumbai University,
Mumbai, India

Chotu Mourya
SCOE, Mumbai University,
Mumbai, India

## ABSTRACT
Recommendation have become extremely common in recent years, and are utilized in a variety of fields, some popular areas include movies, music, news, books, research articles, search queries, social tags, and products in general. They were initially based on demographic, content-based and collaborative filtering. In this project, we are increasing the efficiency rate of recommendation, queried by the user. This is achieved by using an adaptive bandit technique for recommendation- based on exploration-exploitation strategies and classifier technique in multi-armed bandit algorithm. We provide an empirical analysis on medium-size datasets, showing increased prediction performance (as measured by click-through rate).We aim to create recommendation system to predicate with high level of accuracy. We will tackle the cold start problem affecting the system with low amount of user data history.

## General Terms
Knowledge-based system, Explore-Exploitation.

## Keywords
Recommender system, Knowledge-based system, Explore-Exploitation, Artificial intelligence.

## 1. INTRODUCTION
KNOWLEDGE-BASED SYSTEM a computer based system, which uses and generates knowledge from data, information and knowledge. Knowledge based system can act as an expert on demand without wasting time, anytime and anywhere. People use the Internet to express themselves and social networks have become a popular way to share information, ideas and experiences. People use social networks to write sentences with positive, negative emotions, expressing their feelings; in this context, studies concerning sentiment intensity have started to emerge. Knowledge of the sentiment intensity of a sentence can help to collect useful information and allows knowing more about a person who expresses herself or himself about an event, product or content. Sentiment analysis is a technique of natural language processing and text analytics, which can be applied to many areas, such as e-learning, e-commerce, and multimedia among others. There is sentiment analysis research based on physiological signals, subjective emotion assessment , tag-based extractions , web semantic , machine learning, such as, Support Vector Machines (SVM) and its derivation , and the lexicon-based technique such as the ANEW . However, this collection of emotions makes the system dependent on the user input and equipment. It is therefore ideal to have an independent and automatic recommendation system. Tag-based recommendations and lexicon-based techniques Knowledge based system can save money by leveraging Expert, allowing users to function at higher level and promoting consistency.

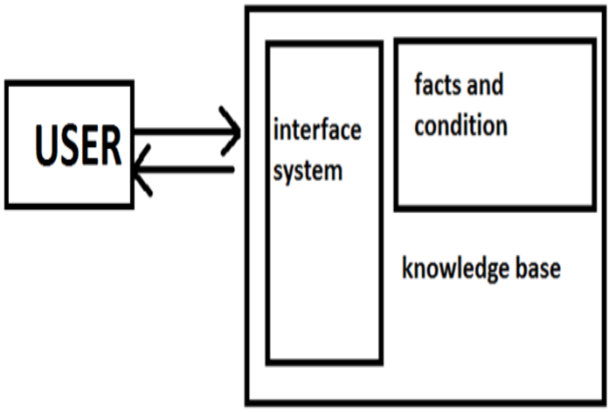


**Fig 1: Knowledge based system**

RECOMMENDATION SYSTEMS are tools for interacting with large and complex information data. They provide a personalized view of such data, prioritizing items likely to be of interest to the user**.** They are the subclass of information filtering system that seeks to predict the "rating" and "preferences" that the user would give to an item. Personalized recommendations are an important part of many on-line e-commerce examples such as Amazon.com, Netflix, and Pandora. The current generation recommendation system is based on collaborative and content-based filtering. The principal disadvantage of collaborative and content based filtering is that systems cannot produce recommendations if there are no ratings available. Secondly, approaches like collaborative filtering and content filtering are plagued by the cold start problem. If there are two words spelled differently but having the same meaning – these filtering will recognize them as two independent words and will not find similarities. Therefore, widely used applications must consume fewer resources in times of an increased number of consumers' portable devices. However, the current generation of recommender systems still requires further improvements to make recommendation methods more effective and applicable to an even broader range of real-life applications. Thus these problems have prompted us to develop new generation .recommendation system based on multi-armed bandit algorithm (epsilon-greedy) and classifier algorithm.

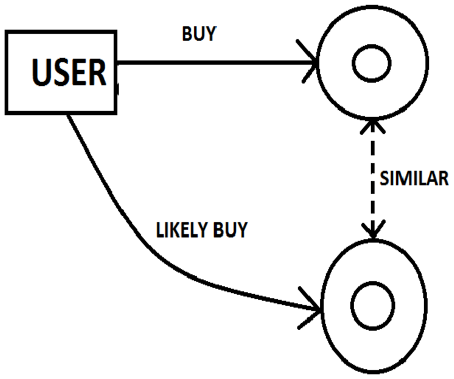


**Fig 2: Recommendation System**

## 2. LITERATURE SURVEY

Recommender Systems (RSs) collect information on the preferences of its users/customer for a set of data (e.g., movies, songs, books, jokes, gadgets, applications, and websites). The get can be acquire explicitly (typically by collecting users' ratings) or implicitly(typically by monitoring users' behavior, such as songs heard, applications downloaded, web sites visited and books read).X demographic features of user may by recommender system like age, nationality. Social information, like twits and posts, is commonly used in websites. Recommender System makes use of different ways of information for providing user/customer with predictions and recommendations of items. Collaborative Filtering (CF) methods play an vital role in the recommendation, although they filtering techniques like content-based filtering or hybrid [1].

Content-based recommendations: The user will be recommended items/data similar to the ones the user preferred in the early process. The content-based has its roots in information retrieval and information filtering research. Content-based systems can recommend text-based items, the content in these systems which is to be recommended is usually described with keywords .In content-based systems, the utility function u(c ,s) is usually defined as: u(c, s) = score(Content Based Profile(c); Content(s)) [2].

Content-based techniques contains limited features that are explicitly associated with the items that these systems recommend. Therefore, in order to have a sufficient set of features, the content must either be in a form that can be parsed automatically by a system (e.g., text) or the features should be assigned to items by user input (manualy) .Another problem with limited content analysis is that, if two different items are in differentiate.

Therefore, since text documents are usually represented by their most important keywords, content-based systems cannot differentiate between a good-written article and a badly written one, if they happen to use the same terms .The user has to rate a sufficient number of items before a content-based recommender system can understand to the user's preferences and provide a reliable recommendation to the user [2]

Unlike content-based recommendation methods, collaborative recommender systems or collaborative filtering systems try to predict the utility of items for a particular based on the items previously rated by other users. The pure collaborative recommender system has some of the shortcomings that content-based systems have. Since collaborative systems use other users' recommendations (ratings), they can deal with any kind of content and recommend any items, even the ones that are not similar to those seen in the past.

However, collaborative systems have their own limitations. New user problem-It is the same problem as with content-based systems. New items are added regularly to recommender systems to reduce the time for recommendation. Collaborative systems rely solely on users' preferences to make recommendations. Therefore, until the new item is rated by number of users, the recommender system would not be able to provide recommendation on it[2].

Several hybrid recommender systems, including Fab and the collaboration based approach are based on traditional collaborative techniques but also maintain the content-based profiles for each user [2]. However there are limitations in hybrid algorithm like the weighted, switching, feature combination and meta-level hybrids were not particularly effective designs for small data set. They showed only scant and spotty improvement over the un-hybridized algorithms, and in the case of meta-level designs, no synergy whatsoever [3].

The multi-armed bandit problem is the problem a gambler faces at a row of slot based machines when deciding which machines will used to play, how many times to play each machine and in which order to play them. When played, each machine provides a random reward from a specific to that machine. The objective of the player is to increase the sum of rewards earned through a lever pulls."Multi-armed bandits come in two flavors: indeterministic and adversarial. The stochastic case is where in response to your actions bandit does not change, while in the adversarial case the bandits learn from user actions and minimize your rewards according to user behaviour. We care about the stochastic case, and our goal is to find the arm which has the greatest *expected outcome*.

I will index the arms by a , and the probability over possible rewards r for each arm a can be written as pa(r). We have to find the arm with the largest mean rewards quickly as possible while collecting the most rewards along the way. One important point is that in practice pa(r)pa(r) are *not stationary* that is, rewards/outcome change over period time and we have to take that into data when we design our algorithms[4].

After studying various existing algorithms like content based, collaborative filtering, hybrid algorithm, multi-armed bandit algorithm, there advantages and disadvantages we came to conclusion that multi-armed bandit algorithms are best for the recommendation system as this algorithm uses best arm(chances) and as soon as possible adjust course when information to the contrary becomes available. This system helps in overcoming suboptimal state forever. The design is essentially to choosing the current best performing arm for a proportion of time. It also guarantees to not get stuck in sub optima state forever. It will use the current best arm for a large time.

The main advantage of these algorithms, it requires low amount of data and they are fast in process with such huge amount of data. It can produce recommendations if there are no ratings available, also changing user preference will not affect the recommendation system .Thus this algorithm will make websites more attractive, catchier providing best.

## 3. PROBLEM STATEMENT

Previously the existing algorithms requires huge amount of data and they are slow in process with such large amount of data. They cannot produce recommendations if there are no ratings available for the certain set of data. Our system aims to develop a learning model for Recommendation search systems based on bandit algorithms.

## 4. PROPOSED SYSTEM

The main aim of project is to develop is fully automated User generated Knowledge based system to recommend over specified data. The algorithm generated is to be helpful for user's for the recommendation purpose over different dataset. The existing algorithm requires huge amount of data & thus they are slow. The algorithm start using best arm(chances). As soon as possible & adjust course when information to the contrary becomes available. This system helps in overcoming suboptimal state forever. The design is essentially to choosing the current best performing arm for a proportion of time.

## 5. FLOWCHART

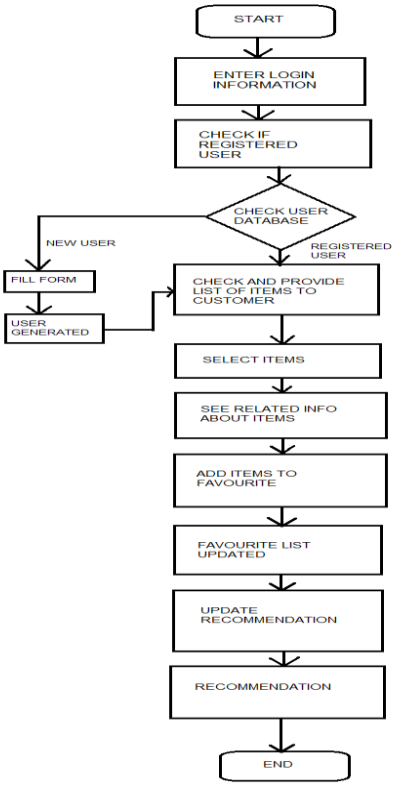


**Fig 3: Flow diagram**

## 6. ALGORITHM (EPSILON-GREEDY)

Epsilon greedy algorithm is a method of solving the bandit problem which deals with the tradeoff between exploitation and exploration .An epsilon value should be chosen between the interval of 0-1.A value closer to 1 will lead to random data which will satisfy the user in the long run but not short term i.e exploration.Whereas exploitation is based on some data set that is present which is used to predict recommendation for the user.

Step 1:-At the start, When a visitor arrives to the site, So the project need to check whether it need to explore or exploit.

Explore:- Choosing at random among all of the arm(options) available to it Exploit:- Choosing the best arm that it currently knows about.

Step 2:- The condition for exploitation& Exploration would generally be,

if Exploit(math.random() > epsilon)

Where, math.random =0 – 1,

Else,

Explore by selecting random arm.

Step 3:- If, The project chooses for exploration an arm will be chosen at random so that the probability of choosing an arm will be equal.

Else, The project follows the exploitation and chose for the best arm which is currently available.

So after selecting the arm a particular reward is given as in, If a user gives positive rating then the reward will be high(i.e 1).

Step 4:-Decrease the Epsilon value so that we it could go from random exploration to exploitation over a period of time(if epsilon value closer to 1 would be standard A/B testing).

Step 5:- After the Explore and exploit function. Plot the readings of feedback for the epsilon value & repeating the Step 2 to Step 4 till epsilon is closer to 0.

Step 6:-Compare feedback number of views or user feedback and obtain Best epsilon value for further recommendation

### 6.1 Psuedocode:-

```
def choose():
ifmath.random() < 0.1:
    # exploration!
    # choose a random lever 10% of the time.
else:
    # exploitation!
    # for each lever,
      # calculate the expectation of reward.
      # This is the number of trials of the lever divided by the total reward
# given by that lever.
   # choose the lever with the greatest expectation of reward.
  # increment the number of times the chosen lever has been played.
  # store test data in redis, choice in session key, etc.
 def reward(choice, amount):
  # add the reward to the total for the given lever.
```

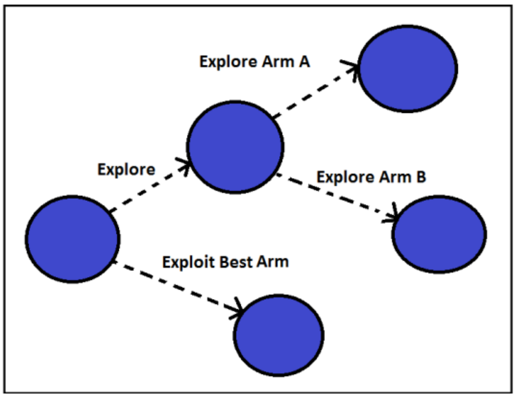


**Fig 4: Explore-Exploit**

## 7. FEATURES OF USER GENERATED RECOMMENDATION SYSTEM USING KNOWLEDGE BASED SYSTEM

1. It guaranteed to not get stuck in sub optima state forever.

2. It will use the current best arm for a large time.

3. The algorithms requires low amount of data and they are fast in process with suchhuge amount of data.

4. It can produce recommendations if there are no ratings available.

5. Items and attributes will be machine-recognizable.

6. Changing user preference will not affect the recommendation system.

7. Time required for data been recommended will decrease by adding multi-armed bandit algorithm which will increase their efficiency of output.

8. This project will used over websites, so that the items get easily recommended to the user from their websites.

## 8. CONCLUSION

The previous Recommendation Systems are tools for interacting with large and complex information space. Thus this project is to develop is fully automated User generated Knowledge based system to recommend over specified data which in turn help over the large sets of data. The latter means that a recommender system does not require any user input to recommend. This system can make an assessment of quality, style or viewpoint by consideration of other people's experience. Items and attributes must be machine-recognizable. Thus this learning model is much efficient, as compared to content-based, collaborative filtering, naïve bayes, etc.

## 9. REFERENCES

[1] J.Babadilla, F.Ortega, A.Hernando. Knowledge-basedsystem. In Elsevier Publication, 2013.

[2] Bobadilla, Jesús, Fernando Ortega, Antonio Hernando, and Abraham Gutiérrez. "Recommender systems survey." *Knowledge-Based Systems* 46 (2013): 109-132.

[3] Kluwer Publication. Hybrid Recommender System Robin Burke, IEEE-2010.

[4] Shauili, AlexandrosKaratzog, Cludio Gentile. Collaborative-Filtering Bandit 30th March 2016.

[5] Liang Tang YexiJiang,Lei li ChunqiuZengTaoLi. Personalized Recommendation via parameter-free Contextual Bandits [ACM-2015].

[6] J´er´emie Mary, Romaric Gaudel, Philippe Preux. "Bandits warm-up cold recommender systems" [University de Lille – July 2014].

[7] John Myles White. "Bandit Algorithms for website optimization"

[8] Djallel Bouneffouf, Amel Bouzeghoub, and Alda Lopes Gançarski. A Contextual-Bandit Algorithm for Mobile Context-Aware Recommender System.

[9] Pavlos Kefalas, Panagiotis Symeonidis, and Yannis Manolopoulos A Graph-Based Taxonomy of Recommendation Algorithms and Systems [IEEE – March 2016].

[10] Renata L. Rosa, Demóstenes Z. Rodríguez, and Graça Bressan Recommendation System Based on User's Sentiments.[IEEE-2015].