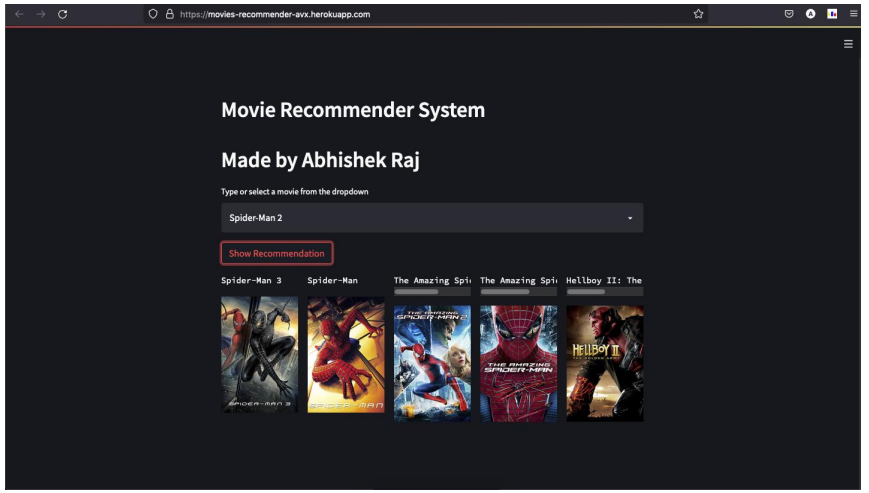
IIIT, Manipur

Movie Recommender System

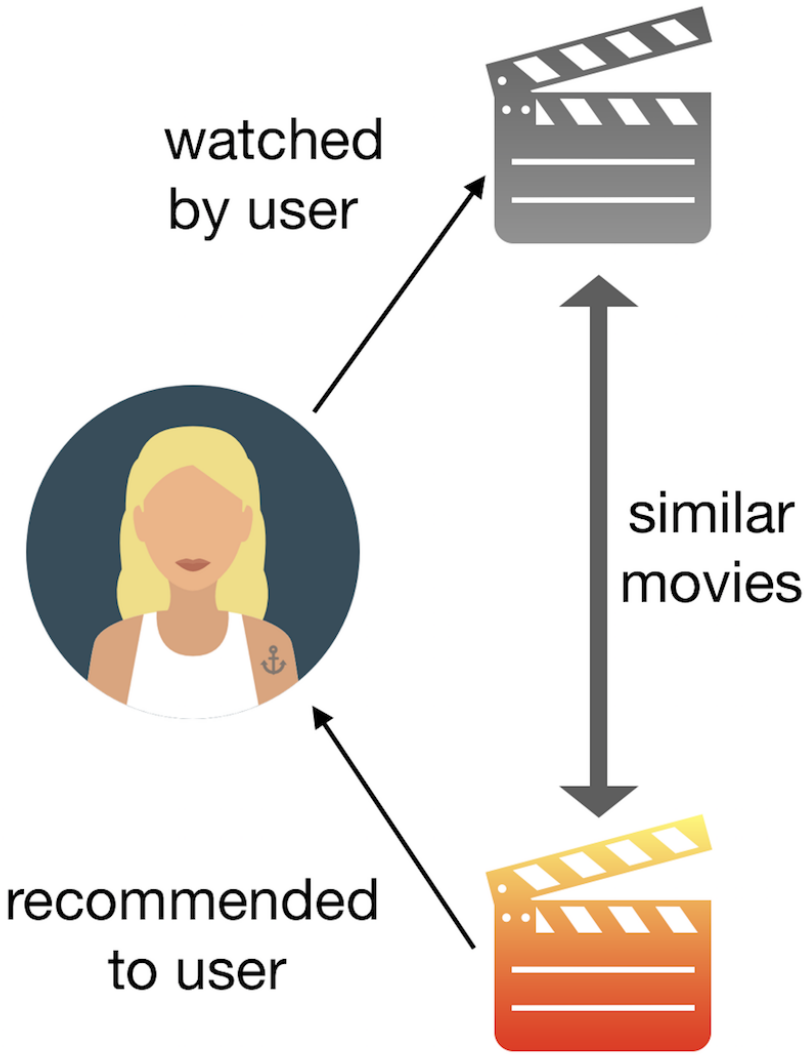


Presented By
Abhishek Raj
Roll No :- 19010112

Guided By
Dr. Kabita Mam
CSE Department

Content

- Abstract
- Theory and fundamental
- Objectives
- Methodology
- Tools and technologies
- Testing / Experimentation
- Conclusion
- Future Enhancement
- References



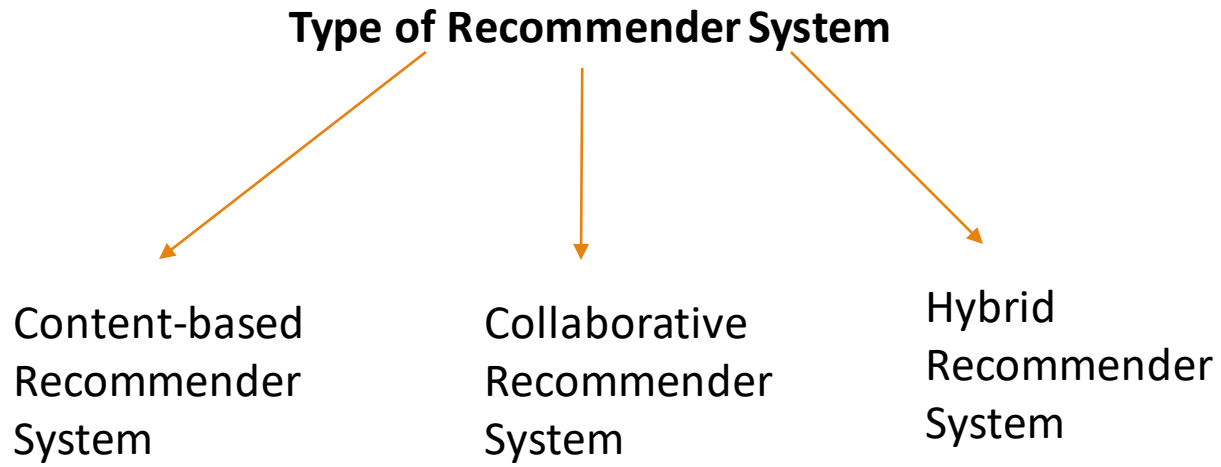
Abstract

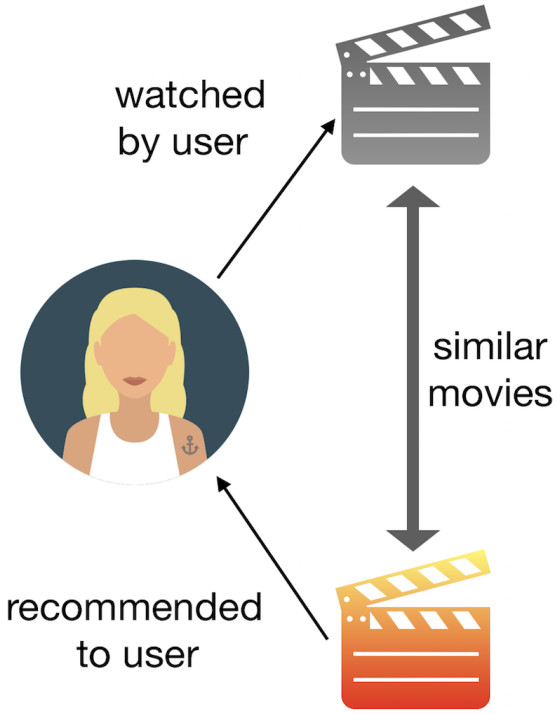
Content-based recommendation systems try to recommend items similar to those a given user has liked in the past. Indeed, the basic process performed by a content-based recommender consists in matching up the attributes of a user profile in which preferences and interests are stored, with the attributes of a content object (item), in order to recommend to the user new interesting items.Content based filtering

Does not require other user's data during recommendations to one user.

Theory and Fundamental

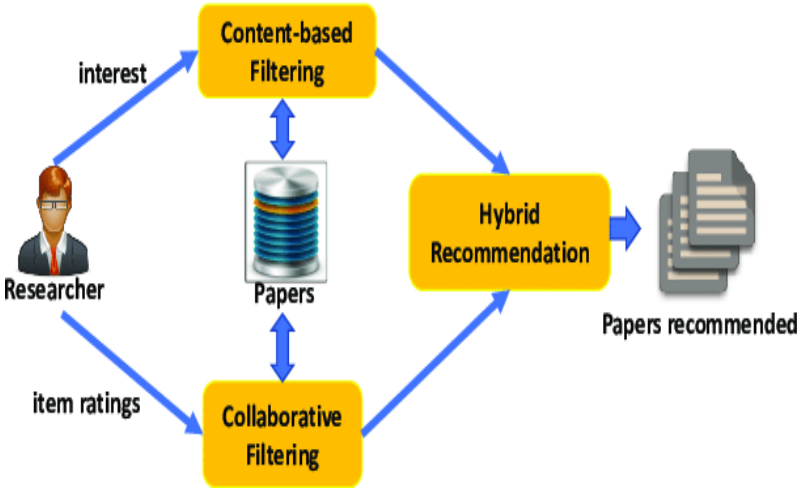
Recommender System :- A recommendation system is a subclass of Information filtering Systems that seeks to predict the rating or the preference a user might give to an item. In simple words, it is an algorithm that suggests relevant items to users.





Content- Based Recommender

Collaborative filtering Recommender System



Hybrid Recommender System

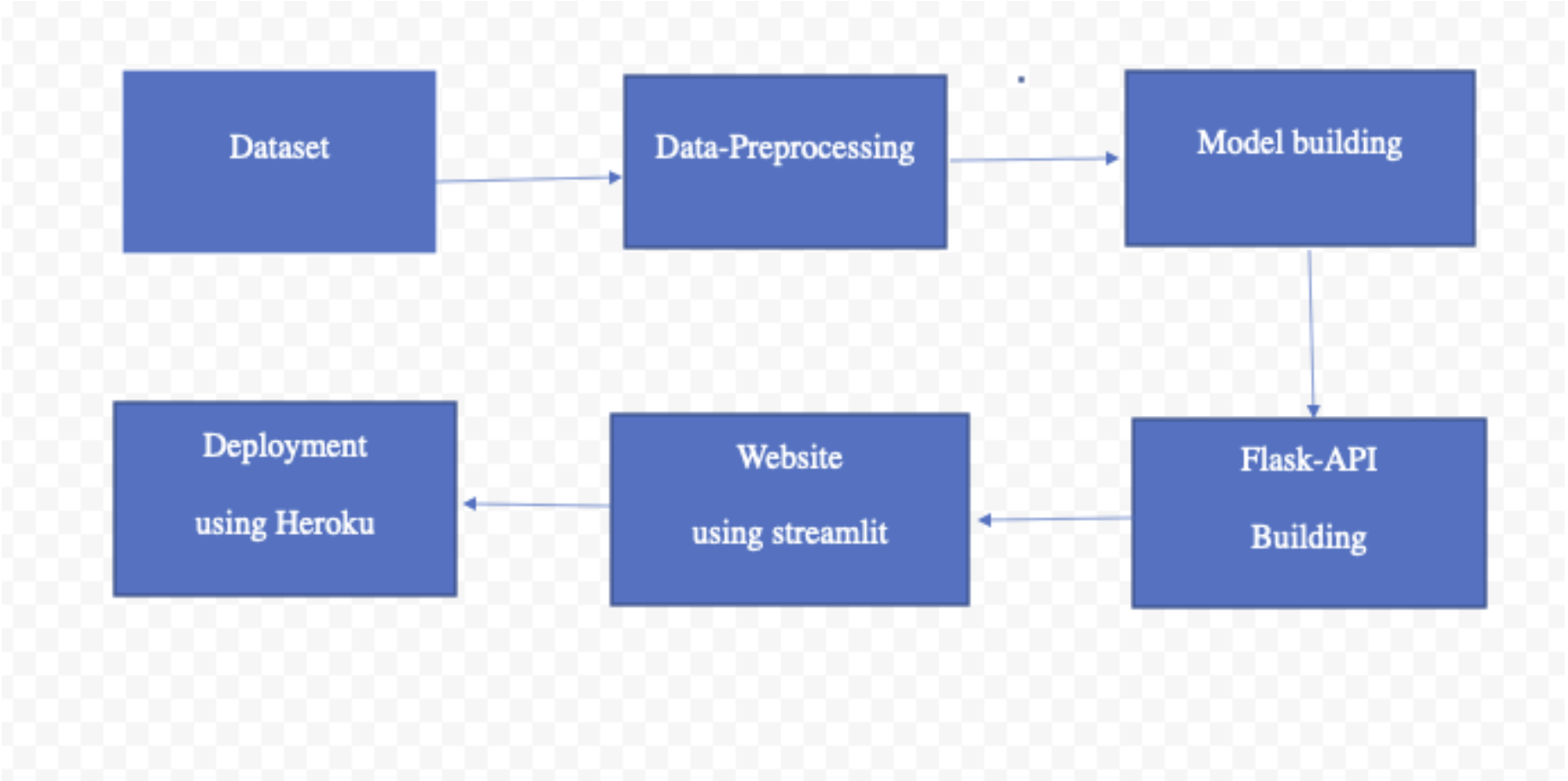
Objectives

- The objective of recommender systems is **to provide recommendations based on recorded information on the users' preferences**. These systems use information filtering techniques to process information and provide the user with potentially more relevant items.

- Here I am going to present **content based recommender System** that uses the **text vectorisation** (text is going to convert into vector) technique like **Bag of word** where we cocatenate all the tags of movies so that we can get most common words by using them we form a **matrix** and by using the **cosine Similarity** technique we can find the similar movie.

- There are many example of Recommender System that uses in real life by Big-gaint company like Netflix , youtube , facebook etc.

Methodology



Flow of Project

1. Dataset :- The first thing we require is data. Data about movies like the **genre, cast, plot** to name a few. For this task, I have taken data from Kaggle. We have two sources of data.

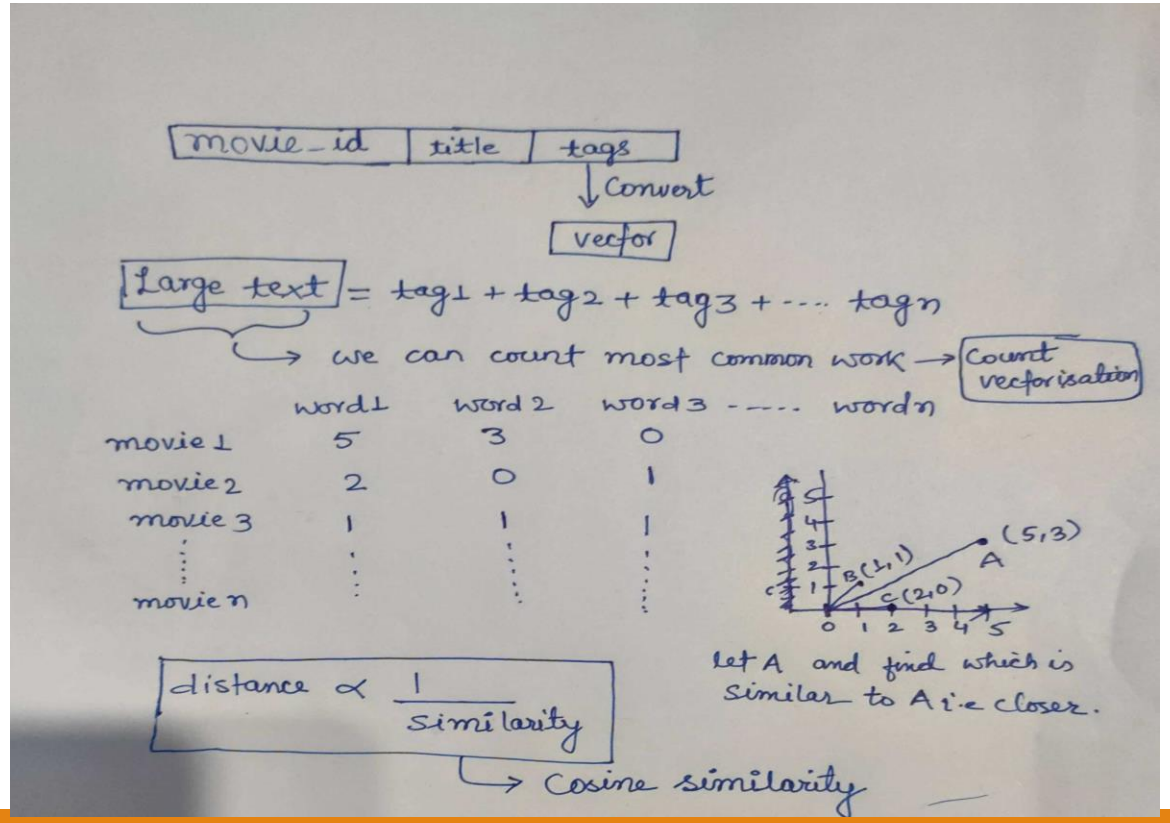
- **TMDB 5000 Movies Data set**
- **The Indian Movie Database**

2. Model Building :- we perform many activity like data preprocessing where we clean the data .
A).**Read** the csv extension dataset.

B). **Combine data** means we have to generate the tags using column like 'overview' , 'genre', ' keyword' , 'cast' and 'crew' after that we drop the column of 'overview' , 'genre' , 'keyword' etc. And form tags form each movie and finally join all the tags of all the movies and a make large sentences so that we can apply text vectorisation technique like **Bags of words.**

Bag of Words is a method to extract features from text documents. These features can be used for training machine learning algorithms. It creates a vocabulary of all the unique words occurring in all the documents in the training set.

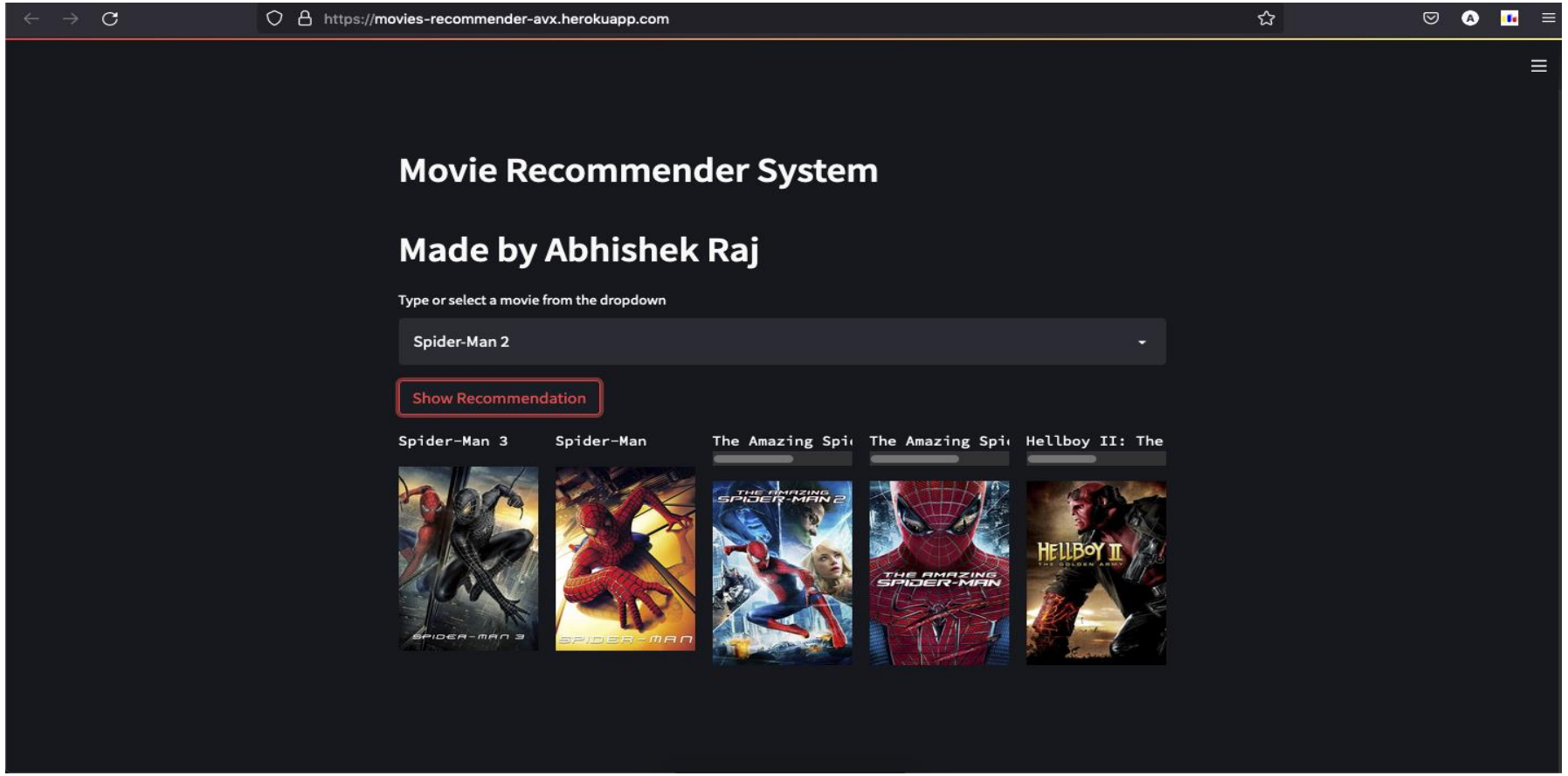
C) Transform data :-



3. Build a website using streamlit :- we fetch the poster of TMDB dataset. Basically we send a request through URL and and get a request in json format which help in recommendation of movie poster along with name.

4. Deployment using Heroku Server.

Testing/ Experiments



Conclusion

Recommender system has become more and more important because of the information overload. For content-based recommender system specifically, we attempt to find a new way to improve the accuracy of the representative of the movie.Recommender systems have their origins in a variety of areas of research, including information retrieval, information filtering, text classification, etc. They use techniques such as machine learning and data mining, alongside a range of concepts including algorithms, collaborative and hybrid approaches, and evaluation methods.

Future work

Recommender system has developed for many years, which ever entered a low point. In the past few years, the development of machine learning, large-scale network and high performance computing is promoting new development in this field. We will consider the following aspects in future work.

- Use collaborative filtering recommendation.
  After getting enough user data, collaborative filtering recommendation will be introduced. As we discussed in Section 2.2, collaborative filtering is based on the social information of users, which will be analyzed in the future research.
- Introduce more precise and proper features of movie
  Typical collaborative filtering recommendation use the rating instead of object features. In the future we should extract features such as color and subtitle from movie which can provide a more accurate description for movie.
- Introduce user dislike movie list.
  The user data is always useful in recommender systems. In the future we will collect more user data and add user dislike movie list. We will input dislike movie list into the recommender system as well and generate scores that will be added to previous result. By this way we can improve the result of recommender system.