Movie Rating Based on users Comments

Amrutha S Nair  
Student  
Department of Computer Science & Engineering  
Mount Zion College of Engineering

Sreelakshmi K  
Assistant Professor  
Department of Computer Science & Engineering  
Mount Zion College of Engineering

Abstract

Movie recommendation system represents the user’s preference for the purpose of suggesting movie. In the proposed system sentiment analysis have been aggregated with a user-based collaborative filtering to provide the accurate recommendation to user. Movie recommendation system proving rating of the reviews on the basis of the reviews of the users, by using sentiment analysis and collaborative filtering techniques. User can select the movie on the basis of categories such as romance, comedy, thriller etc. All users can suggest a movie to others by using rating system.

Keywords: Collaborative filtering, Location sensitive hashing, Recommendation system, Sentiment analysis, negation tagging

I. INTRODUCTION

Recommendation system was originally defined as one in which “people recommendation as input which the system then aggregates and direct to appropriate recipients. Using collaborating filtering it work by searching a large group of users or item finding a smaller list from it with tasks similar to yours. The proposed system is aggregated with collaborative filtering and sentiment analysis. Sentiment analysis is the process of computationally identify and categorizing opinion expressed in a piece of text especially in order to determine whether the writer’s attitude, towards a particular movie. The naive Bayes algorithm is used based on Bayes theorem with an assumption of independence among predictors. The Naive Bayes classifier assumes that the presence of any other features. The proposed frame work the first component is responsible for analysing user reviews and inferring rating from them while the second one is a collaborative filter that generates item recommendation or movie recommendation. To create a review database on the basis of user comments, extract the opinion from the review database and create a opinion database. Compare two databases and take the average to find the rating of movie. All user can suggest a movie to others by using collaborative filtering. Sentiment Analysis is the process of computationally identifying and categorizing opinions expressed in a piece of text.

II. EXISTING SYSTEM

Existing system using location sensitive hashing method. It only focused like and dislike for finding the rating of the movie. It uses randomize algorithms does not provide the guarantee for an exact match. It provides a higher probability of returning the correct answer. It used for reducing the dimension of high-dimension data. It hashes the input data so that same data map to the same bucket with high probability.

A. Limitations of existing system

- It only focused like and dislike for finding the rating of the movie.
- Randomize algorithms does not provide the guarantee for an exact match.
- Higher probability of returning the correct answer.
- High collision approaches.

III. PROPOSED SYSTEM

Proposed system aggregate with collaborative filtering and sentiment analysis. collaborative filtering is used to recommended to other user and sentiment analysis is using uncover the users attitude about a particular movie.

A. Advantages of proposed system

- Calculating both review rating and opinion rating and take the average of the both to finding the total rate of the movie.
- User can recommended to other uses on the basis of rating of the movie.
- The review database creation using two techniques review analysis and data preparation.
- Creating another data base opinion dictionary on the basis of review data base.
- Giving the important to users comments.
– Using sentiment analysis is aim to uncover the attitude of the user on a particular movie from the written text.

### IV. COLLABORATIVE FILTERING

Collaborative filtering is the data based on the collaboration of others. It uses the item matrix in spite of users or item information. User-based collaborative filtering is if uses the matrix to store the ratings of the user. It is a promising technique in recommender systems. It provides personalized recommendations to users based on a database of user preferences, from which users having similar tastes are identified. It then recommends to a target user items liked by other, similar users. It provides personalized recommendations to users based on a database of user preferences, from which users having similar tastes are identified. It then recommends to a target user items liked by other, similar users. CF-based recommender systems can be classified into two major types depending on how they collect user preferences: user-log based and ratings based. User-log based CF obtains user preferences from implicit votes captured through users’ interactions with the system.

#### A. Algorithm

- **Input**: list of users and movies.
- **Output**: similarities of users and movies

```
Begin
$rec ← new Recommended ()
$arc = $rec → get recommendation ( movie, session [id])
for each arc as key → value
    echo $ key ,rating ,$value
$result → select * from visited group by users
while row→ fetch array ($ result )
    echo similarity $session [id] and $ row [user]
$ similarity = $ rec → sim distance ( $ movie ,$ session [id])
End
```

### V. SENTIMENT ANALYSIS

Sentiment analysis is the aim to uncover the attitude of the user on a particular movie from the written text. It is the process of extracting, identifying, analysing, and characterizing the sentiments or opinion in the form of textual information. It identifies the sentiment holders and the entity about which sentiment is expressed. Sentiment analysis is a well-known task in the realm of natural language processing. Given a set of texts; the objective is to determine the polarity of that text. Provides a comprehensive survey of various methods, benchmarks, and resources of sentiment analysis and opinion mining. The sentiments can consist of different classes. A movie review is positive or negative. Where they also employ a novel similarity measure. Sentiment analysis can regroup the opinions of the reviewers and estimate ratings on certain aspects of the product. Another utility of sentiment analysis is for companies that want to know the opinion of customers on their products. They can then improve the aspects that the customers found unsatisfying. Sentiment analysis can also determine which aspects are more important for the customers.

#### A. Algorithm

**Input**: Text file
**Output**: smt = {P,Ng,or N} and strength S, where P:Positive ,Ng: Negative ,N: Neutral

```
Begin
For each ti € T do
    Search for ti in Γ
    If ti € Pos-list then
        SumPos ← SumPos + ti-smt
    Else if ti € Neg-list then
        SumNeg ← SumNeg + ti-smt
    End if
End for
If SumPos >│SumNeg│then
    Smt = P
    S= SumPos / (SumPos + SumNeg)
Else
    Smt = N
    S = SumPos / (SumPos+ SumNeg)
End if
End
```
VI. SYSTEM ARCHITECTURE

Data preparation involves collecting and pre-processing user reviews for the subsequent analysis. Different pre-processing steps may be required depending on the data sources. A user review is likely to be a semi structured document, containing some structured headers and an unstructured text body. Sentiment analysis algorithms usually do not use information other than the comments and the original ratings given by the users. The review analysis step includes several tasks that help identifying interesting information in reviews, which are unstructured, natural language texts. Negation tagging. Some words have negation effects on other words, and negation tagging aims at identifying such words and reflecting their effects when determining the reviews. For example, “good” and “not good” obviously represent opposite sentiments. Feature generalization, or metadata substitution is about generalizing features that may be overly specific. This task can be performed when attributes of domain items are available. For the movie reviews domain, for example, a sentence “Toy Story is pleasant and fun.” in which “Toy Story” is the name of the movie being reviewed, is generalized to “MOVIE is pleasant and fun.” On the basis of review database construct a opinion dictionary. An opinion dictionary contains opinion words, their estimated sentiment oriented and the strengths of their sentiment oriented. Determining the SO and strengths of opinion words is done by answering the question: “Given a certain opinion word, how likely is it to be a positive sentiment, and how likely is it to be a negative one.

VII. CONCLUSION

Most of recommended system can be collect the user details. To create a review database on the basis of user comments. Extract the opinion from the review database and create a opinion database. Compare two databases and take the average to find the rating of movie. All user can suggest a movie to others by using collaborative filtering. Sentiment Analysis is the process of computationally identifying and categorizing opinions expressed in a piece of text.

REFERENCES