Movie Recommender System Using Hybrid Approach For More Accurate Recommendations In Case Of Cold-Start Problem

Mr.A.J.Patankar[1] Abhishar Sinha[2] Anshul Goyal[3] Brajesh Jha[4] Jay Sharma[5] Dr.KotrappaSirbi[6]

Department of Information Technology, D. Y. Patil College of Engineering Akurdi, Pune,[1,2,3,4,5] ,Professor, Computer Science & Engineering, KLE's Dr MSS College of Engg., & Technology, Belagavi,[6]

Abstract— The hybrid approach is one that combines Content Based filtering approach and Collaborative Filtering approach to overcome their short comes and solve the problems mentioned.Pure collaborative filtering approach fails in case of cold start problem. Content based approach can resolve that problem.

1. INTRODUCTION

A recommendation system is a type of information filtering system which attempts to predict the preferences of a user, and make suggests based on these preferences. There are a wide variety of applications for recommendation systems. These have become increasingly popular over the last few years and are now utilized in most online platforms that we use.

2. PROPOSED SYSTEM

The framework we propose comprises of three modules

1. We propose using multiple features rather than using only ratings given by individual users.
2. We propose a hybrid algorithm using both content-based and collaborative filtering methods.
3. We consider taking into account both user and movie similarities.

3. MATHEMATICAL MODEL

Finds similarity between movies based on user ratings matrix.
1. Item Similarity Based Recommendations

User-based recommendation method have been used in many areas, but when the quantity of user data and item data increase fast, problem of data sparseness and scalability of system had been a serious challenge to system. User-based recommendation is also referred to as memory-based recommendation, because the system needs to save all the scoring results to recommend. Therefore, it is possible to building a model to represent the degree of similarity between items, instead of just saving all the rating results.
Show users movies that are most similar to user's liked(high rating) movie.

It uses cosine similarity to calculate similarity.

$$similarity = \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|} = \frac{\sum\limits_{i=1}^{n} A_i B_i}{\sqrt{\sum\limits_{i=1}^{n} A_i^2}\sqrt{\sum\limits_{i=1}^{n} B_i^2}},$$

2. TF-IDF

Term Frequency: the number of times that term t occurs in document d, it is denoted as $tf(t,d) = f_{t,d}$

Inverse Document Frequency: measure of how much information the word provides, i.e., if it's common or rare across all documents.

$$idf(t, D) = \log \frac{N}{|\{d \in D : t \in d\}|}$$

TF-IDF is calculated as

$$tfidf(t, d, D) = tf(t, d) \cdot idf(t, D)$$

A high weight in tf–idf is reached by a high term frequency (in the given document) and a low document frequency of the term in the whole collection of documents; the weights hence tend to filter out common term.

4. SYSTEM SPECIFICATION

Equipment Requirements

•	Speed	: 1.1 GHz.
•	Hard Disk	: 20 GB.
•	Monitor	:15VGA Color.
•	Ram	: 4 GB

Programming Requirements:

•	Operating system	: Any OS
•	Toolkit	: Anaconda
•	Coding Language	: Python 3.6
•	Database	:MongoDB
•	IDE	: PyCharms

International Journal of Research in Advent Technology (IJRAT) Special Issue
E-ISSN: 2321-9637
Available online at www.ijrat.org
National Conference on "Role of Information Technology in Social Innovations"
26ᵗʰ & 27ᵗʰ February 2019
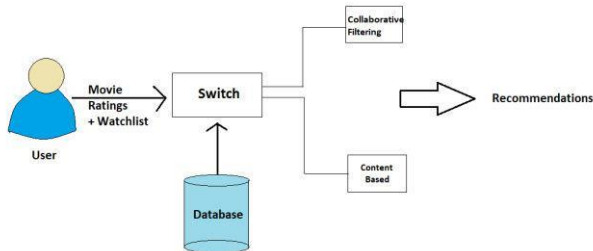
5. SYSTEM ARCHITECTURE

Figure 1: System Architecture of Proposed System

In proposed framework multiple clients can create accounts, login, view recommendations, see movies, give ratings, see ratings, report movies, etc. We also propose producer account who will edit, delete movies and who will be able to see aggregate rating of movies added by him. We also propose movie pages where user can see movie discription and similar movies and will be able to rate it.

User Interface

Figure 2: User Login Page

Figure 3: Recommendation Page

REFERENCES

[1]Jian Wei, Jianhua He, Kai Chen, Yi Zhou, Zuoyin Tang, "Collaborative Filtering and Deep Learning Based Hybrid Recommendation for Cold Start Problem", 2016, IEEE 978-1-5090-4065-0

[2]J. Liu, Y. Jiang, Z. Li, X. Zhang, and H. Lu, "Domain-sensitive recommendation with user-item subgroup analysis," IEEE Trans. Knowl. Data Eng., vol. 28, no. 4, pp. 939–950, Apr. 2016

[3]Shreya Agrawal, Pooja Jain, "An Improved Approach for Movie Recommendation System", IEEE International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)

[4] Chen T, Sun Y, Shi Y, et al. On Sampling Strategies for Neural Network-based Collaborative Filtering[C]//Proceedings of the 23th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM. 2017.

[5]Brovman Y M, Jacob M, Srinivasan N, et al. Optimizing similar item recommendations in a semi-structuredmarketplacetomaximize conversion[C]//Proceedings of the 10th ACM Conference on Recommender Systems. ACM, 2016: 199-202.

[6]Jafarkarimi H, Sim A T H, Saadatdoost R. A naive recommendation model for large databases[J]. International Journal of Information and Education Technology, 2012, 2(3): 216.

[7]P. Resnick, N. Iacovou, etc. "GroupLens: An Open Architecture for Collaborative Filtering of Netnews",

Proceedings of ACM Conference on Computer Supported Cooperative Work, CSCW 1994. pp.175-186

[8] Bijalwan V, Kumar V, Kumari P, et al. KNN based machine learning approach for text and document mining[J]. International Journal of Database Theory and Application, 2014, 7(1): 61-70.