# HYBRID BASED RECOMMENDATION ENGINE: The Art of Matching Items to User

**Divya Vasal[1], Preksha Shukla[2], Vaibhav Vyas[3]**

*[1] Student, AIM & ACT, Banasthali Vidyapeeth, Rajasthan, India*

*[2] Student, AIM & ACT, Banasthali Vidyapeeth, Rajasthan, India*

*[3]Asst. Professor, AIM & ACT, Banasthali Vidyapeeth, Rajasthan, India*

---

**Abstract:** ***The paper focuses on the goals of the recommendation engine which is to generate meaningful recommendations to a collection of users for items or products that might interest them. Recommendation engine have changed the way people find products, information and even other users. It is a study of behavioral patterns for different users to recommend them different items from a collection of things. The design of such recommendation engine depends on the domain and the data available.***

***The traditional approaches are:-***

***1. Content-based filtering-Recommend items similar to the items the user has preferred in the past.***

***2. Collaborative filtering- Recommend items the users with similar tastes preferred in the past.***

***To overcome the limitations of the above mentioned approaches the recent evolving approach is the hybridization of both. A hybrid approach has the potential of delivering enhanced results by exploring the best of both the conventional approaches. Recommendation engine have become trivial tool that helps the developers to generate an algorithms for the prediction of items which a user may prefer among the other list of given items. In recent years, they have been applied in a variety of applications like Pandora, NETFLIX, YouTube, FACEBOOK, LinkedIn etc. The proposed Hybrid Approach provides improvements in addressing two major challenges of recommendation engine: accuracy of the engine and sparsity of data by simultaneously incorporating correlation between items and users. The evaluation of the system shows superiority of the solution compared to stand-alone collaborative and content based approaches. This can be considered very important, especially when dealing with a large quantity and diversity of resources.***

**Key Words:** *Recommendation engine ,Content-Based Filtering, Collaborative Filtering, and Hybrid Approach.*

## 1. INTRODUCTION

With the overloading of information and increase in the number of variety of products, a new need in technology emerged -RECOMMENDATION ENGINE.

Recommendation engine is a subclass of information filtering system that seek to predict the 'rating' or ' preference' that user would give to an item. They provide personalized suggestions to the user. Manyon-line stores provide recommending services; example- **AMAZON, IMDB, NETFLIX, YouTube, FACEBOOK, LinkedIn** and so forth.

Three basic steps for a conventional recommendation engine [4]:-
An input is provided by the user to the engine. These inputs can be stated clearly or implied. The ratings given by the user for a particular item are readily observable whereas the various websites that the user visits that is the click-through rate are inherited for the future recommendations.

The user likes and dislikes are based on these ratings and their click rate value. This representation could be as simple as a model of items-ratings, or as complex as a data structure combining both content and ratings.
Then suggestions are made to the user based on their profile.

The above steps are same for most of the recommendation engines, but there are many approaches present like: COLLABORATIVE FILTERING and CONTENT-BASED APPROACH are basic approaches while HYBRID
FILTERING is a new approach to improve the quality of the
Recommendations.

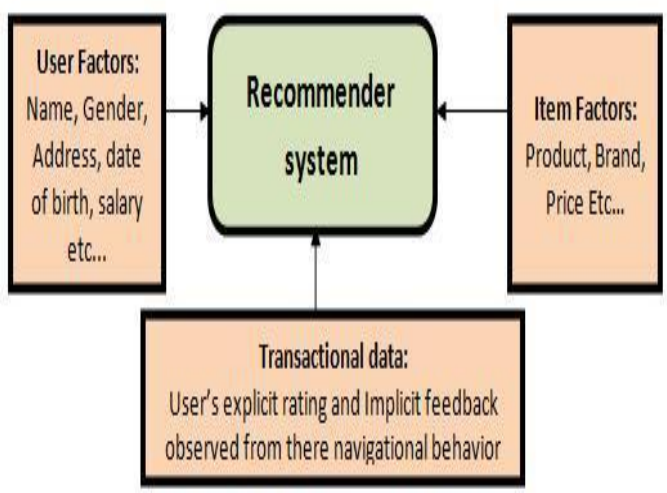


**Fig 1:** Recommendation engine [3]

## 2. LITERATURE REVIEW

A variety of techniques have been proposed as the basis for recommender engine: Collaborative Filtering and Content based Approach have been surveyed below:

### 2.1 Collaborative based Recommendation systems

The items are recommended to a user by pairing his/her personal taste with the other users. Collaborative systems apply the nearest neighbor model for computing recommendations. This model consists of three steps:

The users provide ranks to the products they have experienced before.

Likes and dislikes of the active and other users are matched. To do so, correlation coefficients are calculated between them.

The users whose score highly correlate with that of the active user are called the neighbors.
Pearson Correlation is a standard method for computing correlations.

The collaborative filtering technique recommends items based on user-based approach and item-based approach. [3]

1) User-Based Approach:
In this approach the user plays an important role. If certain majority of the customer has same taste then they form one group. Recommendations are given to user based on the common preferences he/she shares in the same group. If the item was positively rated by the community, it will be recommended to the user.

2) Item-Based Approach:
Here the items play an important role Recommendations are based on ratings of items. The system develops recommendations on the basis of items in the neighborhood that a user would prefer.

Amazon recommender engine is based on collaborative filtering which uses item based collaborative filtering (people who buy y also buy z). Other examples include: Facebook, Twitter, LinkedIn, and other social networking sites use collaborative filtering to recommend new friends (based on the connections of the user).

### 2.2 Content Based Recommendation System:

The system recommends items based on its given content rather than the other users ratings. There are typically four steps for content based recommendations: [4]

Gather appropriate data for various items. Example: beats, singer, lyrics etc. for the music or the story, cast etc.for the movies. Most systems use Information

Retrieval techniques to retrieve the relevant information.

In this step, the user is asked to rate random items he/she prefers.

The user profile is recapitulated using the content information derived in the first step and the ratings provided in the second step.

To match unrated item contents with the user profile gathered in the third step and assigning scores to the items depending on the match. The items are ranked according to their scores and presented to the user in that particular order.

**Pandora Radio** is a prominent example of a content-based recommender engine that plays music with similar traits to that of a song provided by the user as an initial seed.

All of the basic approaches (collaborative, content-based) suffer from some shortcomings such as:

**Cold Start**: These approaches require a large amount of existing data on a user in order to make accurate recommendations.

**Scalability**: Due to the enormous number of users and items a large amount of computation is required to calculate recommendations.

**Sparsity**: The number of items bought by the set of users on major e-commerce sites is very large. The most active users will only have rated a small subset of the overall database. Thus, even the popular items have low ratings.

**Gray Sheep**: The user does not consistently agree or disagree to the group of people and due to this reason user recommendation seems to be difficult. [3]

**Explainability**: It is another important aspect of recommender systems. An incomplete reasoning such as "you will like this item because you liked those items" due to this recommendations of items is difficult. [3]
Hybrid recommender systems are those that combine many techniques such as the ones described above to improve recommendation performance and achieve some synergy between them.

## 3. METHODOLOGY

The term **Hybrid Recommender System** describes a recommender system that combines multiple approaches together to produce an output. Many recommendation engine use a **hybrid** approach by combining collaborative and content-based methods, which helps to avoid certain limitations of these approaches.

There are different types of Hybrid Systems: [1]

- **Weighted**: It is the score of different recommendation components which are combined numerically. *The example of weighted method is P-TANGO system which uses hybrid Recommendations. Here equal weights are assigned to both content and collaborative recommenders but gradually weights are adjusted as the prediction of ratings is confirmed*

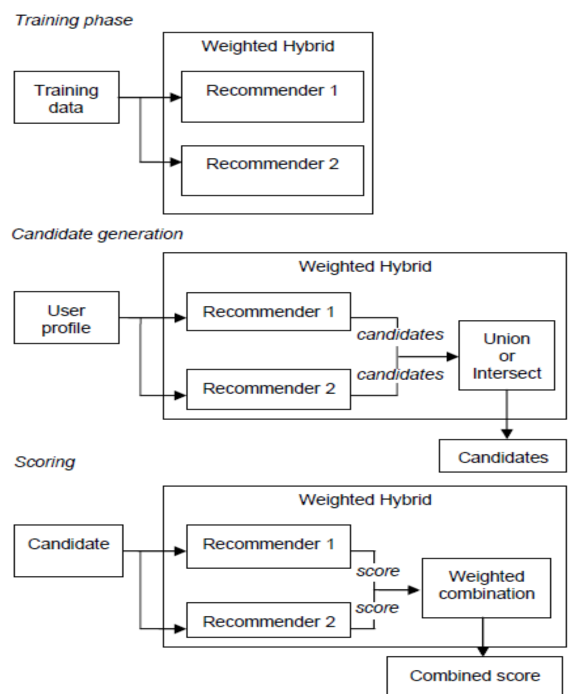


**Fig 2:** Weighted Hybrid [1]

- **Switching**: The system chooses among recommendation components and applies the selected one. *The DailyLearner system uses a combination of content and collaborative approaches in which a content-based recommendation method is applied first. If the content-based system does not give accurate* *recommendation, then a collaborative recommendation is attempted.*

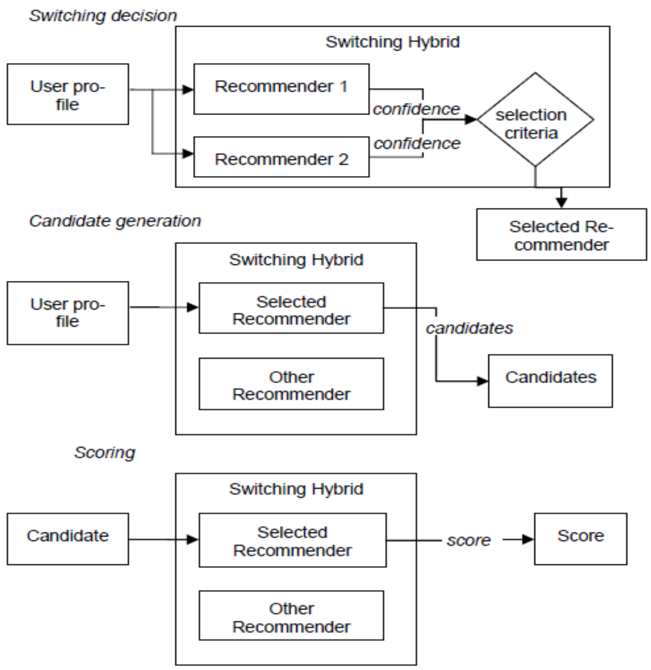


**Fig 3:** Switching Hybrid [1]

• **Mixed**: Recommendations from different approaches are presented together. *This method is used in Television System. Content based method is used for textual description of TV-shows and use of collaborative filtering is for finding the preferences of the user and recommendations from these two methods suggest a final program.*

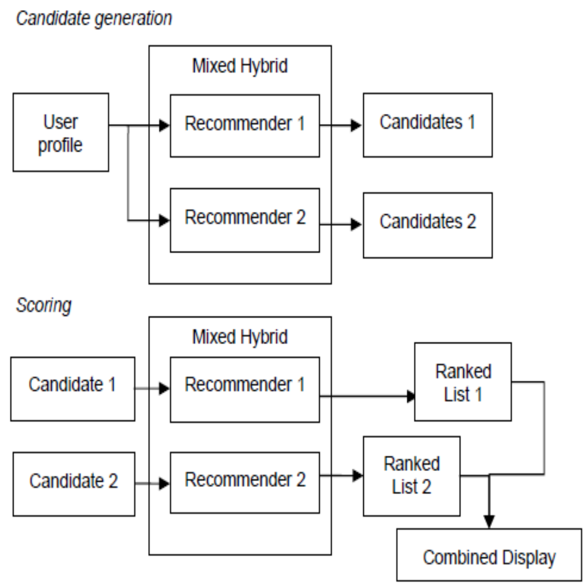


**Fig 4*:*** Mixed Hybrid [1]

• **Cascade**: This is basically the refinement of different techniques where one technique is give higher priority than the other. *The Entree restaurant recommender was found to return too many equally-scored items, which could not be ranked relative to each other. Then the hybrid EntreeC was created by adding a collaborative re-ranking of only those items with equal scores.*

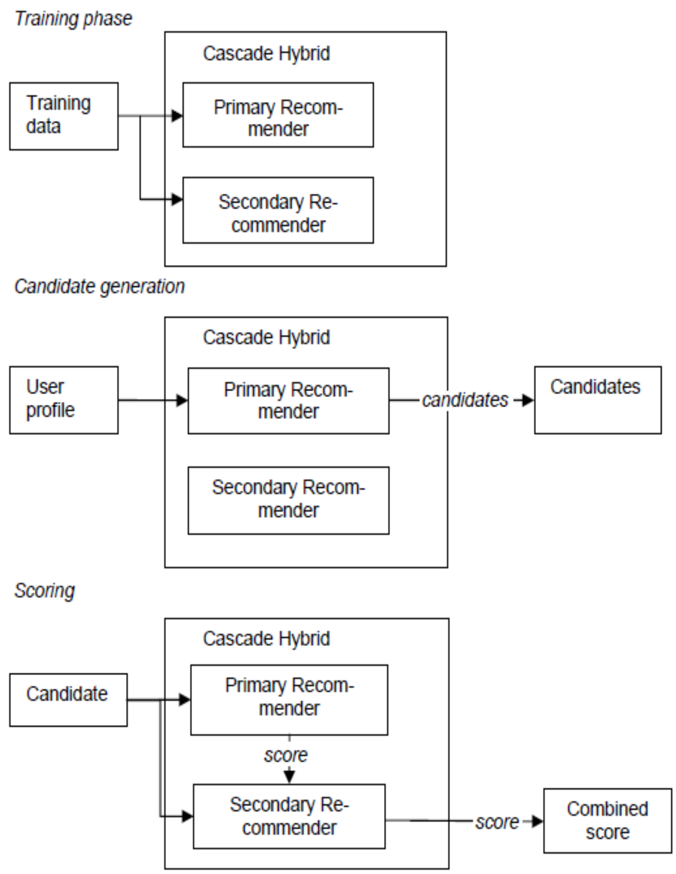


**Fig 5:** Cascade Hybrid [1]

• **Meta-level**: One recommendation technique is applied and produces some sort of model, which is then the input used by the next technique.*Pazzani used the term "collaboration through content" to refer to his restaurant recommender that used to build models of user preferences in a content-based way. With each user so represented, a collaborative step was then being performed in which the peer users were identified.*

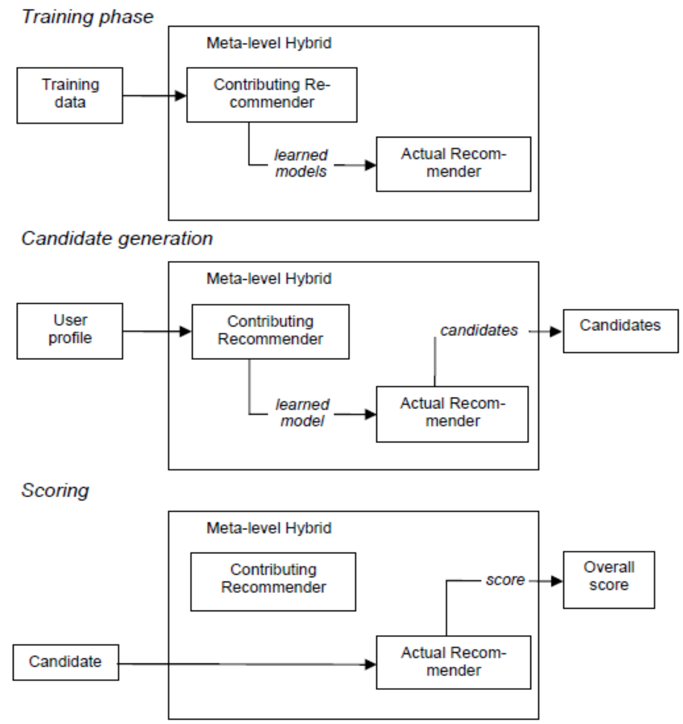


**Fig6:** Meta-level Hybrid [1]

•**Feature Augmentation**: One recommendation technique is used to compute a feature or set of features, which acts as an input to the next technique. *Melville, Mooney and Nagarajan invented the term "content-boosted collaborative filtering." This algorithm learns a content-based model over the training data and then uses this model to induced ratings for unrated items. This makes for a set of profiles that is compact and more suitable to the collaborative stage of recommendation that does the actual commending.*

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 02 Issue: 07 | Oct-2015 www.irjet.net p-ISSN: 2395-0072

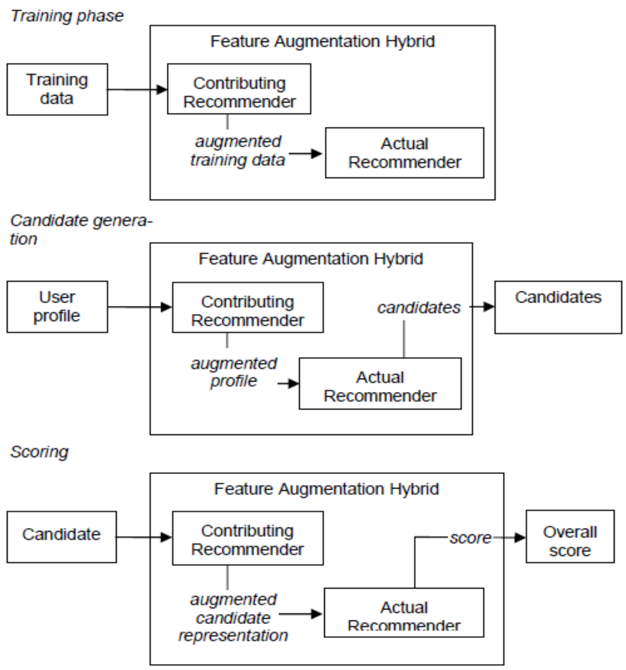


**Fig 7:** Feature Augmentation Hybrid [1]

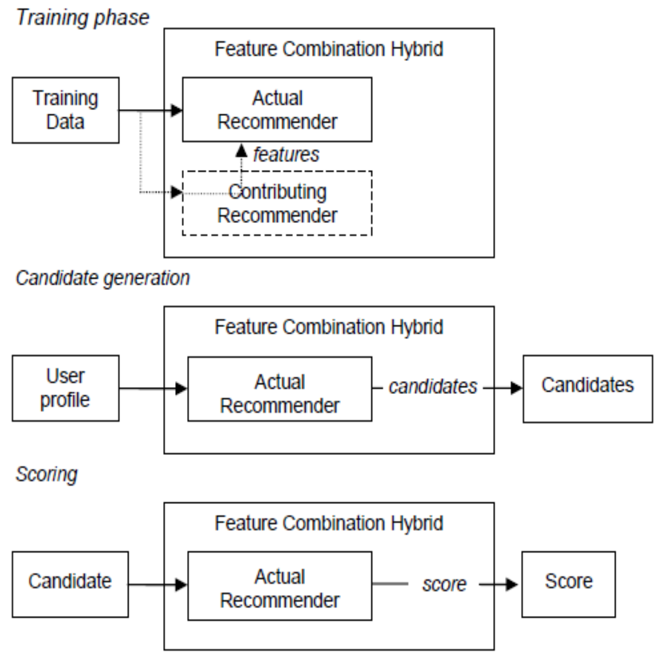


**Fig 8:** Feature Combination Hybrid [1]

•**Feature Combination**: Features deduced from various sources of information are integrated together and given to a single recommendation algorithm.*Basu, Hirsh and Cohen learnt the content-based rules about user's likes and dislikes from the inductive rule learner Ripper. They improved the system's performance by summating the collaborative features, thereby evaluating a fact like "User A and User B liked a show" in the same way that the algorithm treated features like "Actor X and Actor Y starred in the show".*

Several recommendation systems use a *hybrid* approach. Different ways to combine collaborative and content-based methods into a hybrid recommender system can be classified as follows: [2]

1. **Combining separate recommenders.** In this method to implement the hybrid recommendations we have to implement separate collaborative and content methods. Then after, we can unify the ratings gathered from individual recommender engine into one ultimate recommendation using either a linear combination of ratings or a voting scheme.
Example: DailyLearner System: selects the recommender system that can give the recommendation with the higher level of confidence, while chooses the one whose recommendation is more consistent with past ratings of the user.

2. **Adding content-based characteristics to collaborative models.** This type is used to overcome the problem of Sparsity. In this approach the similarity between two users is calculated between the content based profiles and not commonly rated items. Another benefit of this approach is that users can be recommended an item not only when this item is rated highly by users with similar profiles, but also directly,

that is, when this item scores highly against the users profile.

**3. Adding collaborative characteristics to content-based models.** The most popular approach is to use some dimensionality reduction technique on a group of content-based profiles and creates collaborative view over a collection of users' profiles.

**4. Developing a single unifying recommendation model.** This approach is getting very popular among researchers. For instance, one approach proposes to use characteristics of both content-based and collaborative approaches (Example: age or sex of users or the kind of movies). To sum up, it uses the profile information of users and items in a single statistical model that estimates unknown ratings for other users and items.

Through these techniques many drawbacks of collaborating and content based approaches were overcome and improved the understanding between users and items.

## 4. CONCLUSION

Since 1990s recommendation engine have been very important and useful tool for filtering relevant information, in many fields such as: education, entertainment, restaurant and tourism activities. The literature of recommendation engine however shows that individual recommender engines such as: collaborative filtering and content-based filtering have some drawbacks such as: sparsity, cold start problems and other relevant problems. Hybrid recommendation engine combines collaborative and content-based approaches to solve individual problems of these techniques. Although Recommendation engine made a meaningful progress over the last decennary on several content-based, collaborative filtering and hybrid approaches. Despite all the advancement, the present generation of recommendation engine assessed in this paper still stand in need for some further enhancements to make recommendation approaches more effective in a wide range of applications. In this paper, various limitations of the current recommendation engines are discussed and methods are provided to overcome these problems. Thus we conclude that integrating collaborating filtering and content based techniques can significantly improve predictions of a recommender engine.

## 5. REFERENCES

[1] Robin Burke, *Hybrid Recommender Systems: Survey and Experiments*, School of Computer Science, Telecommunications and Information Systems, DePaul University, 243 S. Wabash Ave. Chicago, Illinois, USA,2002.

[2] Gediminas Adomavicius and Alexander Tuzhilin, *Towards the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions,* IEEE Transactions on knowledge and Data Engineering, Piscataway, NJ, USA: pp.734-749, 6June 2005, vol.17.

[3] Jaimeel Shah and Lokesh Sahu, *A Survey of Various Hybrid based Recommendation Method*, Parul Institute of Engg & Tech, Vadodara, Gujarat, India, November 2014, vol.4.

[4] Mustafa Bilgic,*Explanation for Recommender Systems Satisfaction vs. Promotion,* Department of Computer Sciences, University of Texas, Austin, May 19,2004.