



Project Presentation on BOOK RECOMMENDATION SYSTEM USING CONTENT BASED (TAG BASED) FILTERING

**Presented by:
Deeksha Arora (50385)
Maansi Bhatnagar (50387)
Rakshita Joshi (50388)**



CONTENTS

1. Introduction
2. Motivation
3. Classification of Recommendation System
4. Problem Statement
5. Requirements
6. Methodology
7. Results
8. Conclusion
9. References

INTRODUCTION

- Recommendation system is an information filtering technique, which provides users with information, which he/she may be interested in.
- It aimed at suggesting relevant items to users (items being movies to watch, text to read, products to buy or anything else depending on industries).
- EXAMPLES:-

NETFLIX

Video-on-demand provider in North America and UK

- Matches 23 million customers with a huge inventory of movies according to their tastes
- 60-70% of views result from the recommendations⁶

amazon.com

Gold standard of e-commerce. Pioneer in using recommendations

- Sits on a huge volume of collective information of its customers
- Customers can view what people with similar tastes viewed or purchased
- Customers can ask the recommendations engine to ignore selected purchases

Linked in
facebook

Social and professional networking sites

- Sits on a huge volume of collective information of its customers
- Customers can view what people with similar tastes viewed or purchased
- Customers can ask the recommendations engine to ignore selected purchases

PANDORA

Music station. Offers music suggestions based on ratings

- Sits on a huge volume of collective information of its customers
- Customers can view what people with similar tastes viewed or purchased
- Customers can ask the recommendations engine to ignore selected subscriptions³

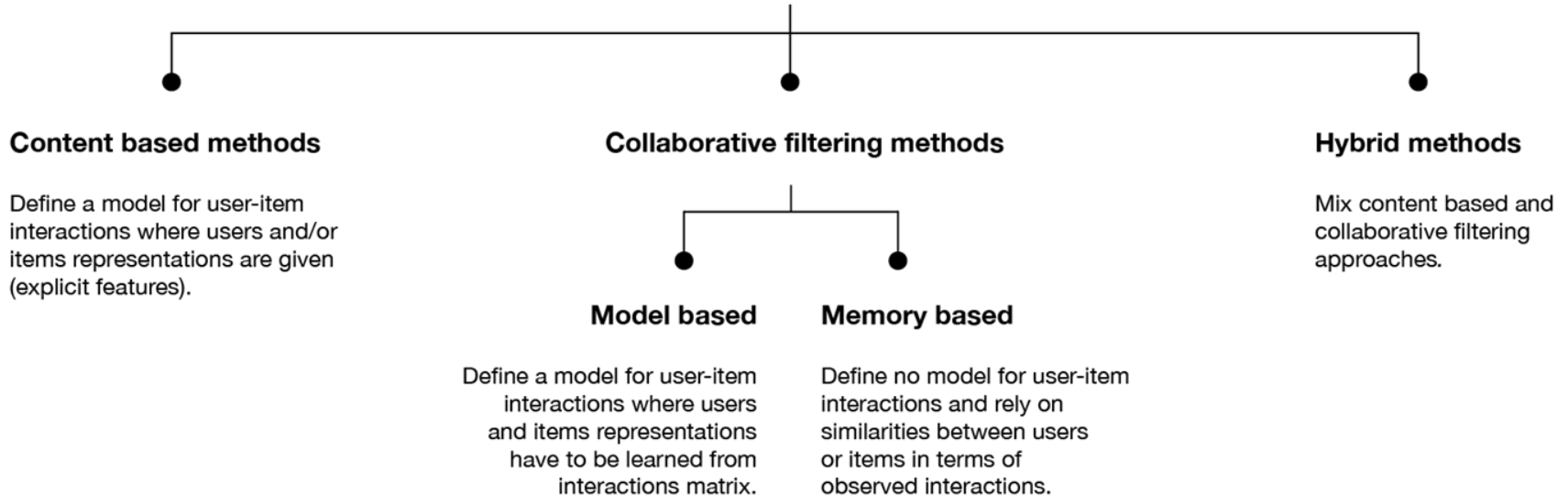


MOTIVATION

- help the user to get personalized recommendations
- helps users to take correct decisions in their online transactions
- increase sales and redefine the users web browsing experience
- retain the customers
- enhance their shopping experience.

Classification of Recommendation System

Recommender systems



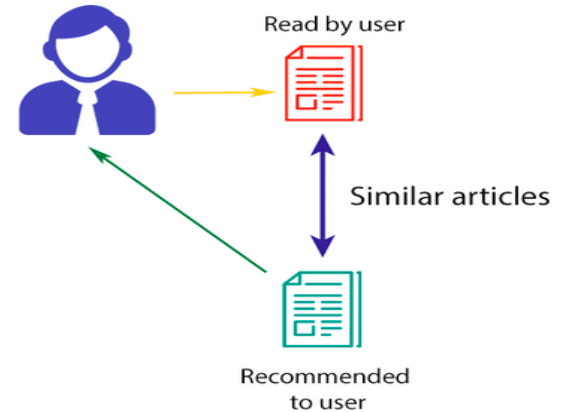
CONTENT BASED FILTERING

- recommend items similar to those a user has liked (browsed/purchased) in the past.

OR

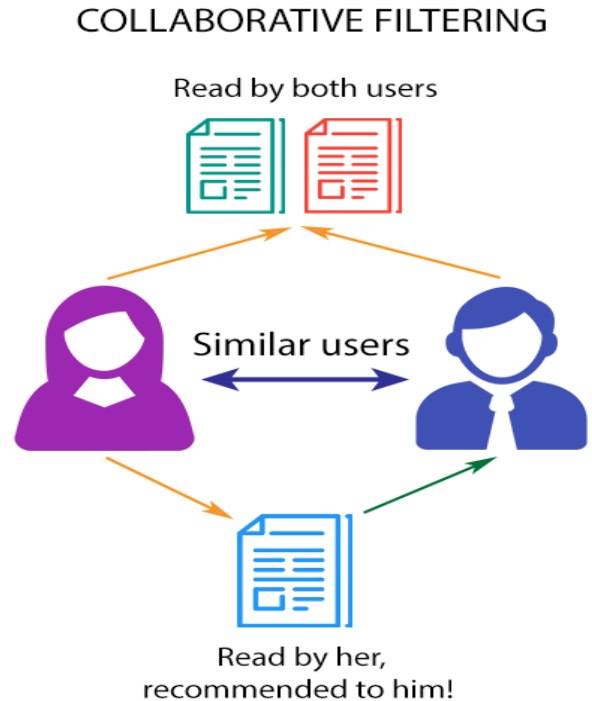
- Recommendations are based on the content of items rather on other user's opinion

CONTENT-BASED FILTERING



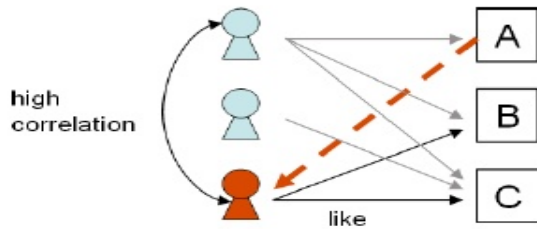
COLLABORATIVE FILTERING

- finds a subset of users who have similar tastes and preferences to the target user and use this subset for offering recommendations
- Main Approaches :
 - Memory based
 - Model based



MEMORY BASED

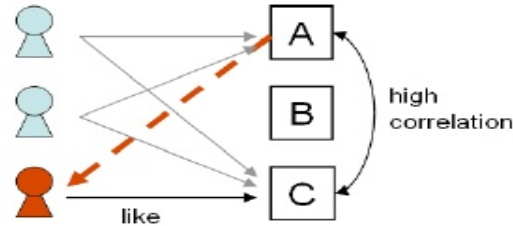
- The main characteristic of this approach is that they use information only from the user-item interaction matrix and they assume **no model** to produce new recommendations.



User-based filtering
(Grouplens, 1994)

Take about **20-50** people who share **similar taste** with you, afterwards predict how much you might like an item depended on how much the others liked it.

You may like it because your "friends" liked it.



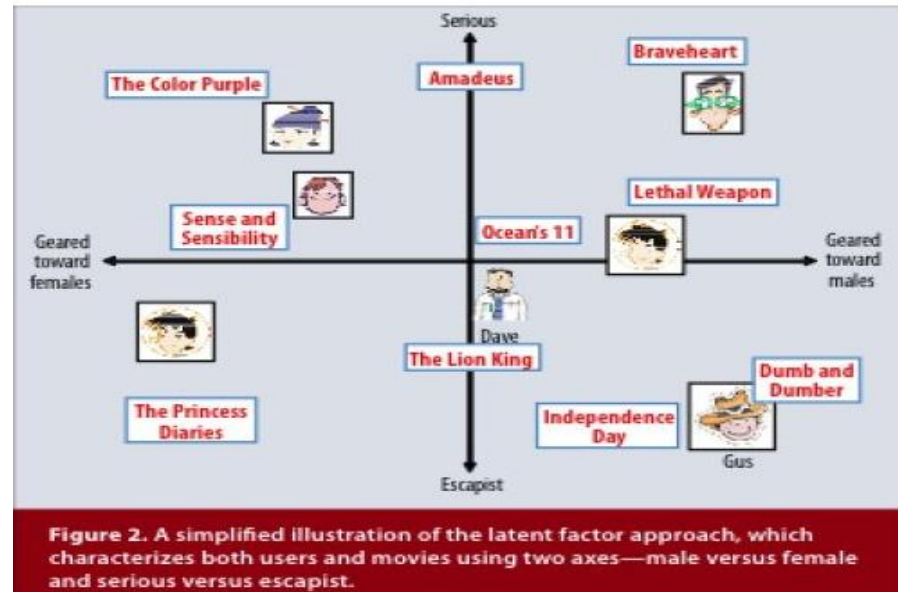
Item-based filtering
(Amazon, 2001)

Pick from your previous list **20-50** items that share **similar people** with "the target item", how much you will like the target item depends on how much the others liked those earlier items.

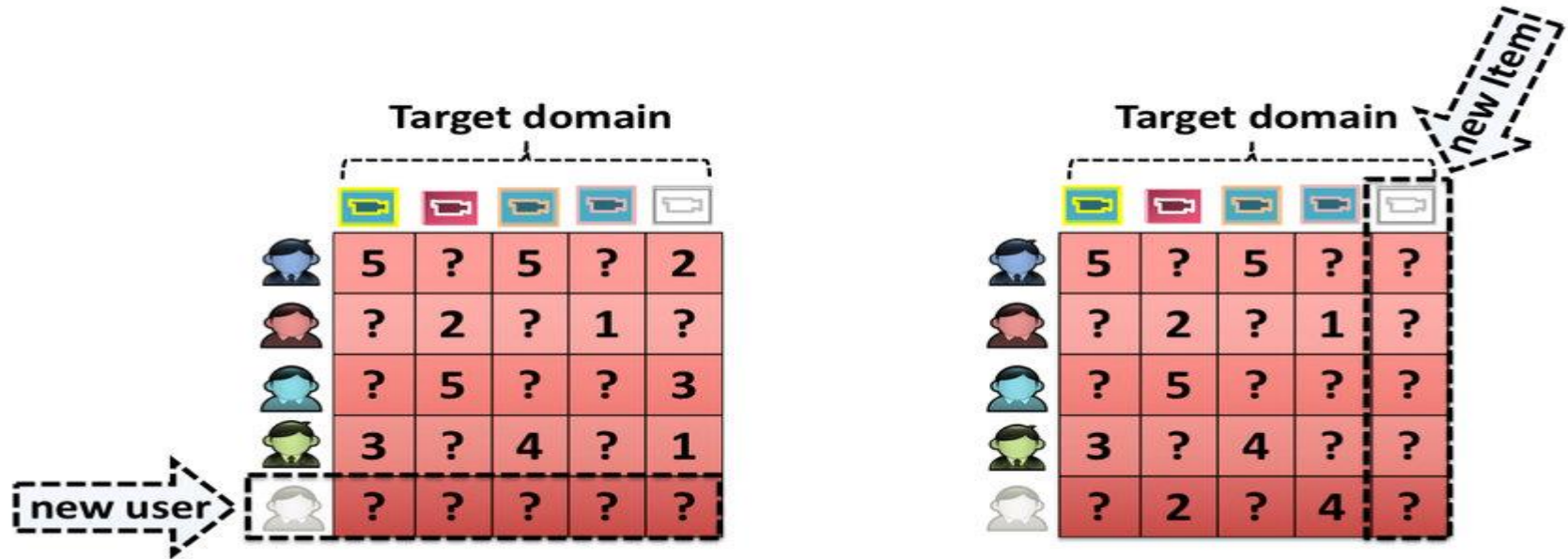
You tend to like that item because you have liked those items.

MODEL BASED

- only rely on user-item interaction information and assume a latent model supposed to explain these interactions



COLD START PROBLEM



SOLUTION :- RANDOM STRATEGY

PROBLEM STATEMENT

Given some datasets (books, ratings, tags, booktags, to read) we develop A Book Recommendation System using:-

Content Based Filtering (Tag Based)



REQUIREMENTS

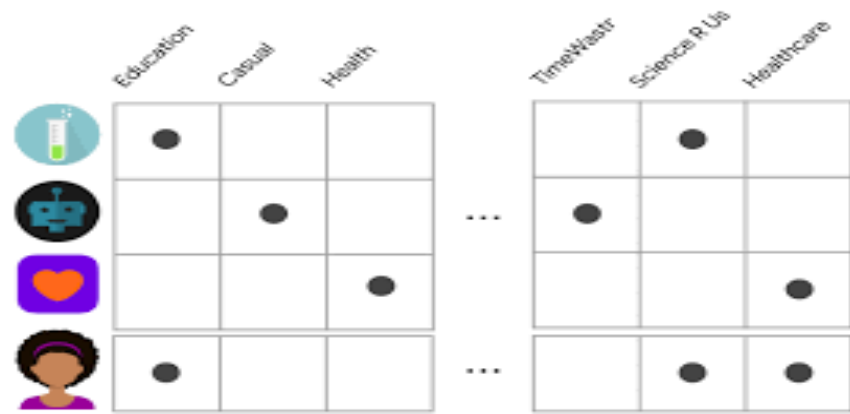
- **Software Interfaces**
 1. Microsoft Word 2016
 2. Database - My SQL 5.6
 3. Operating System: Windows10
 4. Anaconda 3
- **Hardware Requirements:**
 1. i5 Processor Based Computer
 2. 4 GB-Ram(Atleast)
 3. 100 GB Hard Disk(Atleast)



CONTENT BASED MODEL

UNDERSTANDING THE CONCEPT

- Content-based filtering uses item features to recommend other items similar to what the user likes, based on their previous actions or explicit feedback.
- Some of the user-related features could be explicitly provided by the user. For example, a user selects "Entertainment apps" in their profile.



The diagram illustrates content-based filtering by showing two user profiles and their interactions with various app categories. Each profile is represented by a vertical stack of icons and a corresponding grid of interest points (black dots).

Profile 1 (Left): Represented by icons for a test tube, a robot, a heart, and a person. The grid shows interest in Education, Casual, and Health categories.

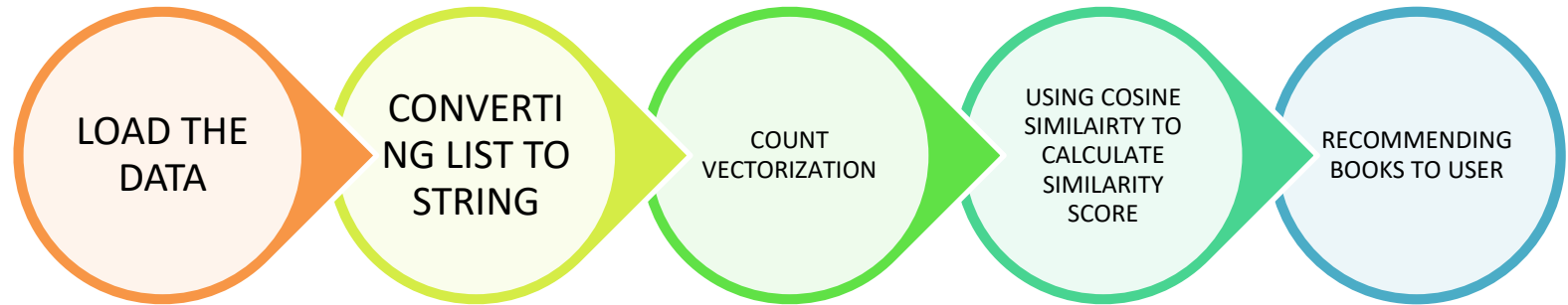
	Education	Casual	Health
Test Tube	●		
Robot		●	
Heart			●
Person	●		

Profile 2 (Right): Represented by icons for a test tube, a robot, a heart, and a person. The grid shows interest in TimeWastr, Science P.Us, and Healthcare categories.

	TimeWastr	Science P.Us	Healthcare
Test Tube		●	
Robot	●		
Heart			●
Person		●	●

Ellipses (...) indicate that there are more categories and users in the system.

STEPS INVOLVED



CONVERTING LIST TO STRING

- In order to be able to properly perform content based filtering, first we have to put all the tags related to one single book in a single string: each tag must be separated from each other with a space, to be accounted as a single word.

COUNT VECTORIZATION

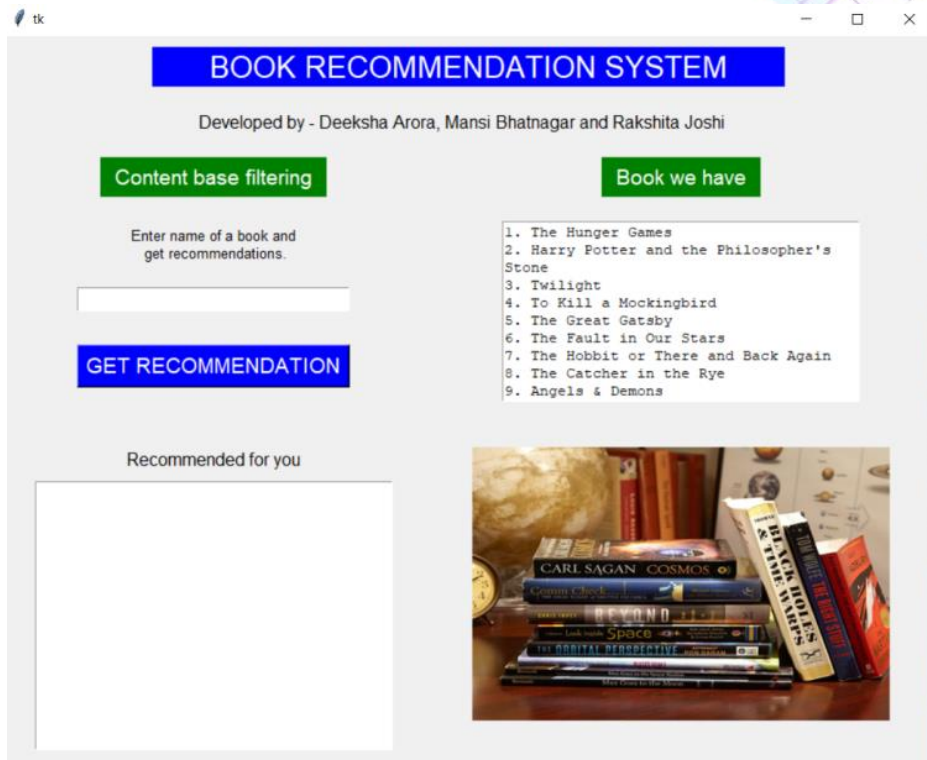
- In order to use textual data for Content based modeling, the text must be parsed to remove certain words – this process is called **tokenization**. These words need to then be encoded as integers, or floating-point values, for use as inputs in machine learning algorithms. This process is called **feature extraction (or vectorization)**.
- CountVectorizer is used to convert a collection of text documents to a vector of term/token counts.
- It will generate a matrix in which the columns represent all the tag-words that have been included in the analysis and the rows represent the books.

COSINE SIMILARITY

- Cosine similarity is a metric used to determine how similar the documents are irrespective of their size.
- Mathematically, it measures the cosine of the angle between two vectors projected in a multi-dimensional space.

$$\text{similarity}(A,B) = \frac{A \cdot B}{\|A\| \times \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n A_i^2} \times \sqrt{\sum_{i=1}^n B_i^2}}$$

RESULTS



User visits the Desktop app where he/she can select a book of his/her choice from a list of books.

tk

BOOK RECOMMENDATION SYSTEM

Developed by - Deeksha Arora, Mansi Bhatnagar and Rakshita Joshi

Content base filtering


Enter name of a book and get recommendations.

GET RECOMMENDATION

Book we have

1. The Hunger Games
2. Harry Potter and the Philosopher's Stone
3. Twilight
4. To Kill a Mockingbird
5. The Great Gatsby
6. The Fault in Our Stars
7. The Hobbit or There and Back Again
8. The Catcher in the Rye
9. Angels & Demons

Recommended for you



User enters the book of his choice and Clicks on "Get Recommendation" button

BOOK RECOMMENDATION SYSTEM

Developed by - Deeksha Arora, Mansi Bhatnagar and Rakshita Joshi

Content base filtering

Enter name of a book and
get recommendations.

GET RECOMMENDATION

Recommended for you

1. The Hunger Games
2. Harry Potter and the Philosopher's Stone
3. Twilight
4. To Kill a Mockingbird
5. The Great Gatsby
6. The Fault in Our Stars
7. The Hobbit or There and Back Again
8. The Catcher in the Rye
9. Angels & Demons
10. Pride and Prejudice

Book we have

1. The Hunger Games
2. Harry Potter and the Philosopher's Stone
3. Twilight
4. To Kill a Mockingbird
5. The Great Gatsby
6. The Fault in Our Stars
7. The Hobbit or There and Back Again
8. The Catcher in the Rye
9. Angels & Demons



Books are
recommended
to the user.

CONCLUSION

The goal of most of the recommendation systems is to predict the reader's interest and recommend the books accordingly. This book recommendation has considered many parameters like content of the book, author of the book and quality of the book by using content filtering of ratings by other readers. Our experience while working on this project has enhanced our skills and has contributed towards our professional and personal development.

REFERENCES

1. https://www.researchgate.net/publication/325792825_An_Improved_Online_Book_Recommender_System_using_Collaborative_Filtering_Algorithm
2. <http://infolab.stanford.edu/~ullman/mmds/ch9.pdf>
3. Stuart E. Middleton, Nigel R. Shadbolt (2004). Ontological user profiling in recommender systems. ACM Transactions on Information Systems, 54-88
4. <http://taste.sourceforge.net/>

A decorative geometric pattern in the top-left corner, consisting of overlapping squares and rectangles in various shades of blue and purple, creating a complex, layered effect.

THANK YOU