Target

# Recommendation System

9th November 2020

## CASE STUDY

Given the list of books and ratings for books.Implement a recommender system to suggest Books to:

- existing users,
- new users,
- friends of existing users

Implement a module that takes a user name and returns recommendations (5 books in specific order of recommendation). Showcase score indicating the success of your recommender system.

Describe the approach of scoring the recommender system's success. How would you updated your system and react to when a user does one of the below:

- accepts a recommendation from 5 suggested
- If a user rejects all your recommendations?

## GOALS

1. Exploratory Data Analysis
2. Overview of Book Reviews
3. Personalization of Users based on their likes and dislikes
4. Recommendation System for New User
5. Recommendation System for Existing User
6. Recommendation System for Friends of Existing User
7. Recommendation Matrix Update

## Tech Stack

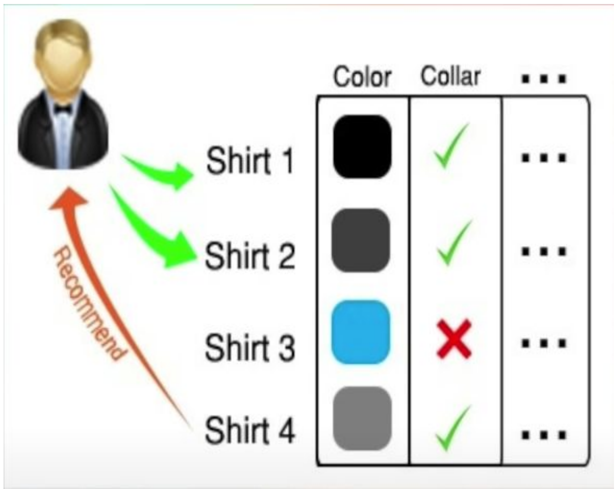
1. Python
2. Jupyter lab
3. Python Libraries like Pandas, Numpy, Scipy
4. ML Models: KNN, LightFM, TFRS
5. Deep Learning Framework: Tensorflow, TFRS (Latest release in sept 2020)

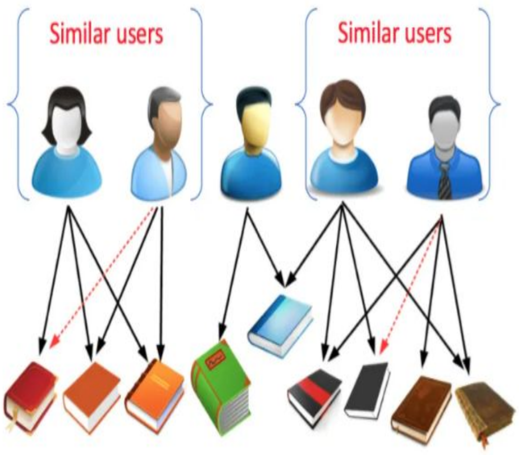
## APPROACH:

Recommendation Systems are of two types:

1. Content Based Filtering
2. Collaborative Based Filtering

There are ways to use both the Content based and Collaborative Based Filtering. My plan was to build an Ensemble Collaborative Based Filtering with Content Based Filtering.





### My Approach

1. Cluster Similar Books and Similar Users
2. In a group of Similar Users, User's liked books which are not read by other similar users. For eg: Person A,B,C having the same likes. If A has not read a book which B and C together or individually liked the book.
3. Content Based Filtering
4. Collaborative Based Filtering
5. Ensemble User Preference Recommendation
6. Hybrid Approach

## MILESTONES

### Exploratory Data Analysis

1. Data Preparation
2. Data Cleaning:
    a. Null Checks
    b. Duplicate Users
3. User Level Exploratory Data Analysis
4. Book Level Exploratory Data Analysis
5. Clustering Users and Books

Google Sheet for data:
https://docs.google.com/spreadsheets/d/1ogYRCQgxJZTSHSbVPWSdxN-kA1DxOpRLMlCDeFjZW0g/edit#gid=0

### Content Based Filtering

This is based on a single user's interactions and preference. Recommendations are based on the metadata collected from a user's history and interactions.

I have used Google Books API to get features like Genre, Number of Pages and many other metadata.

### Clustering

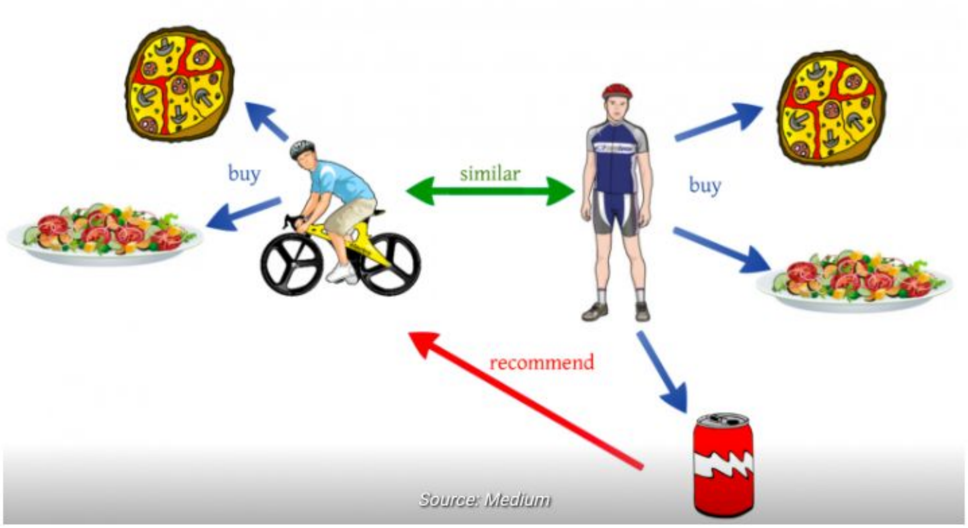
Group the user and book with Similar Ratings. Cluster Similar users and similar books.

### Collaborative Based Filtering

Collaborative filtering casts a much wider net, collecting information from the interactions from many other users to derive suggestions for you. This approach makes recommendations based on other users with similar tastes or situations.

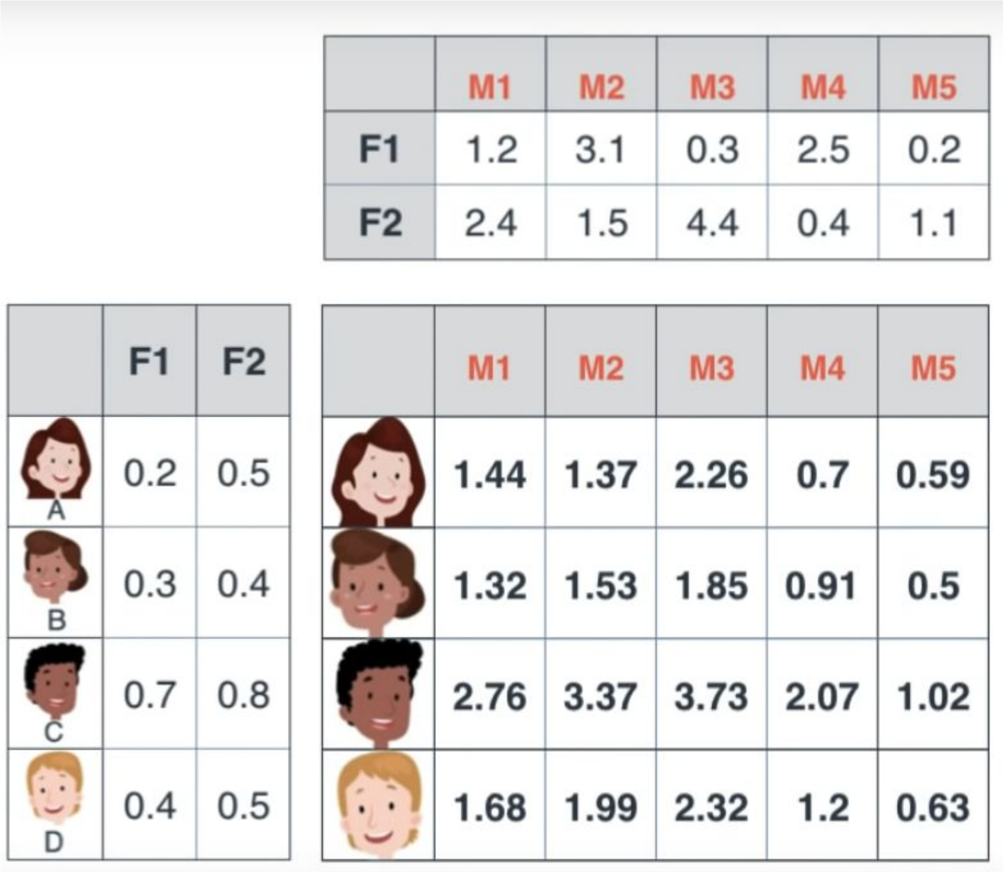
#### User - User Collaborative Filtering:

This algorithm first finds the similarity score between users. Based on this similarity score, it then picks out the most similar users and recommends products, which these similar users have liked or bought previously.



Source: Medium

The method **Matrix Factorization** could be of great help to identify if the user will like the book based on his/her preference.



| | M1 | M2 | M3 | M4 | M5 |
|---|---|---|---|---|---|
| F1 | 1.2 | 3.1 | 0.3 | 2.5 | 0.2 |
| F2 | 2.4 | 1.5 | 4.4 | 0.4 | 1.1 |

| | F1 | F2 |
|---|---|---|
| A | 0.2 | 0.5 |
| B | 0.3 | 0.4 |
| C | 0.7 | 0.8 |
| D | 0.4 | 0.5 |

| | M1 | M2 | M3 | M4 | M5 |
|---|---|---|---|---|---|
| | 1.44 | 1.37 | 2.26 | 0.7 | 0.59 |
| | 1.32 | 1.53 | 1.85 | 0.91 | 0.5 |
| | 2.76 | 3.37 | 3.73 | 2.07 | 1.02 |
| | 1.68 | 1.99 | 2.32 | 1.2 | 0.63 |

Item

$$\text{User}\;\begin{array}{c|cccc} & W & X & Y & Z \\ \hline A & & 4.5 & 2.0 & \\ B & 4.0 & & 3.5 & \\ C & & 5.0 & & 2.0 \\ D & & 3.5 & 4.0 & 1.0 \end{array} = \begin{array}{c|cc} A & 1.2 & 0.8 \\ B & 1.4 & 0.9 \\ C & 1.5 & 1.0 \\ D & 1.2 & 0.8 \end{array} \times \begin{array}{cccc} W & X & Y & Z \\ \hline 1.5 & 1.2 & 1.0 & 0.8 \\ 1.7 & 0.6 & 1.1 & 0.4 \end{array}$$

Rating Matrix

User Matrix

Item Matrix

|  | M1 | M2 | M3 | M4 | M5 |
|---|---|---|---|---|---|
| F1 | 3 | 1 | 1 | 3 | 1 |
| F2 | 1 | 2 | 4 | 1 | 3 |

|  | F1 | F2 |
|---|---|---|
| A | 1 | 0 |
| B | 0 | 1 |
| C | 1 | 0 |
| D | 1 | 1 |

|  | M1 | M2 | M3 | M4 | M5 |
|---|---|---|---|---|---|
| A | 3 |  | 1 |  | 1 |
| B | 1 |  | 4 | 1 |  |
| C | 3 | 1 |  | 3 | 1 |
| D |  | 3 |  | 4 | 4 |