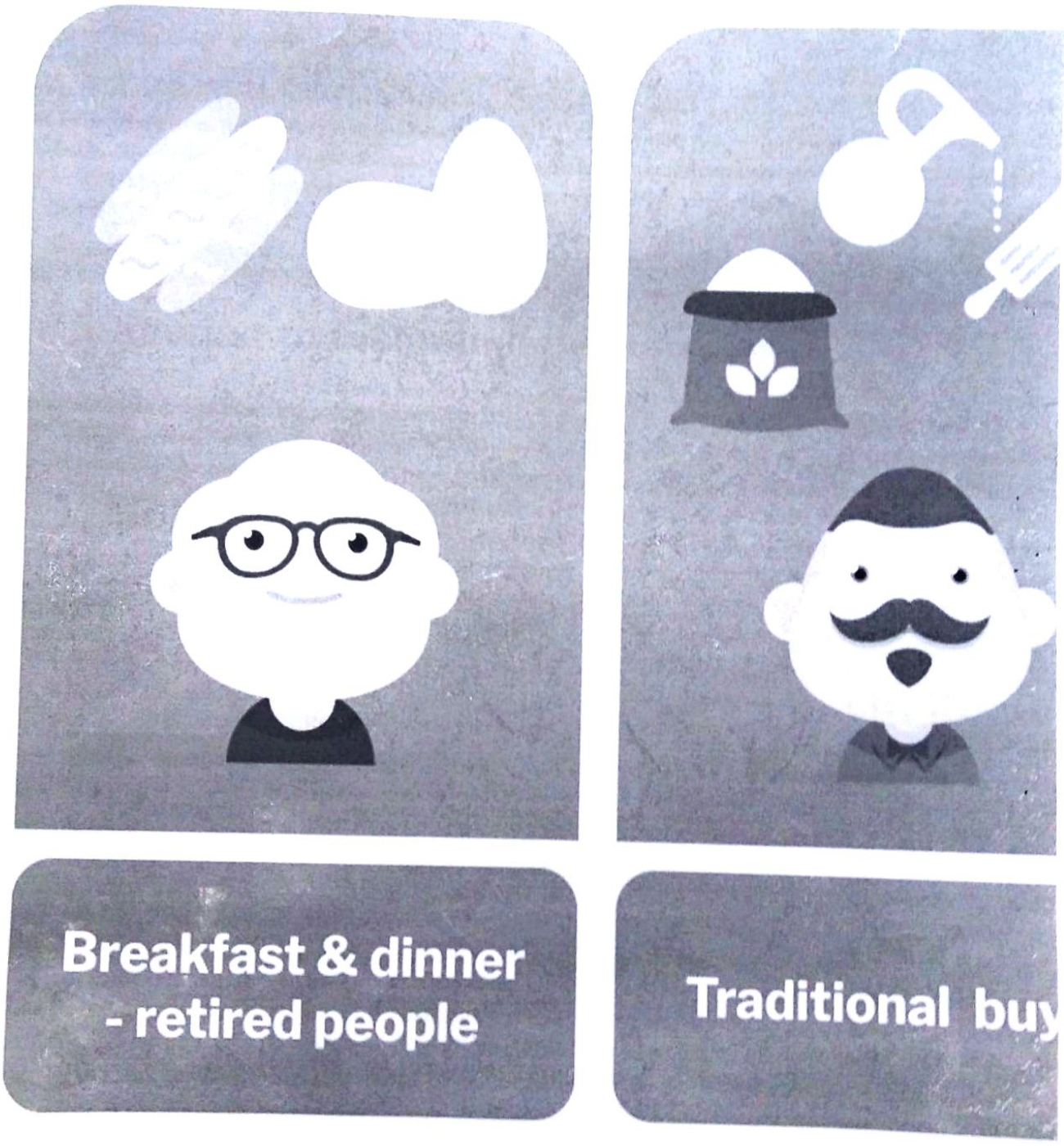
Introduction to recommender systems

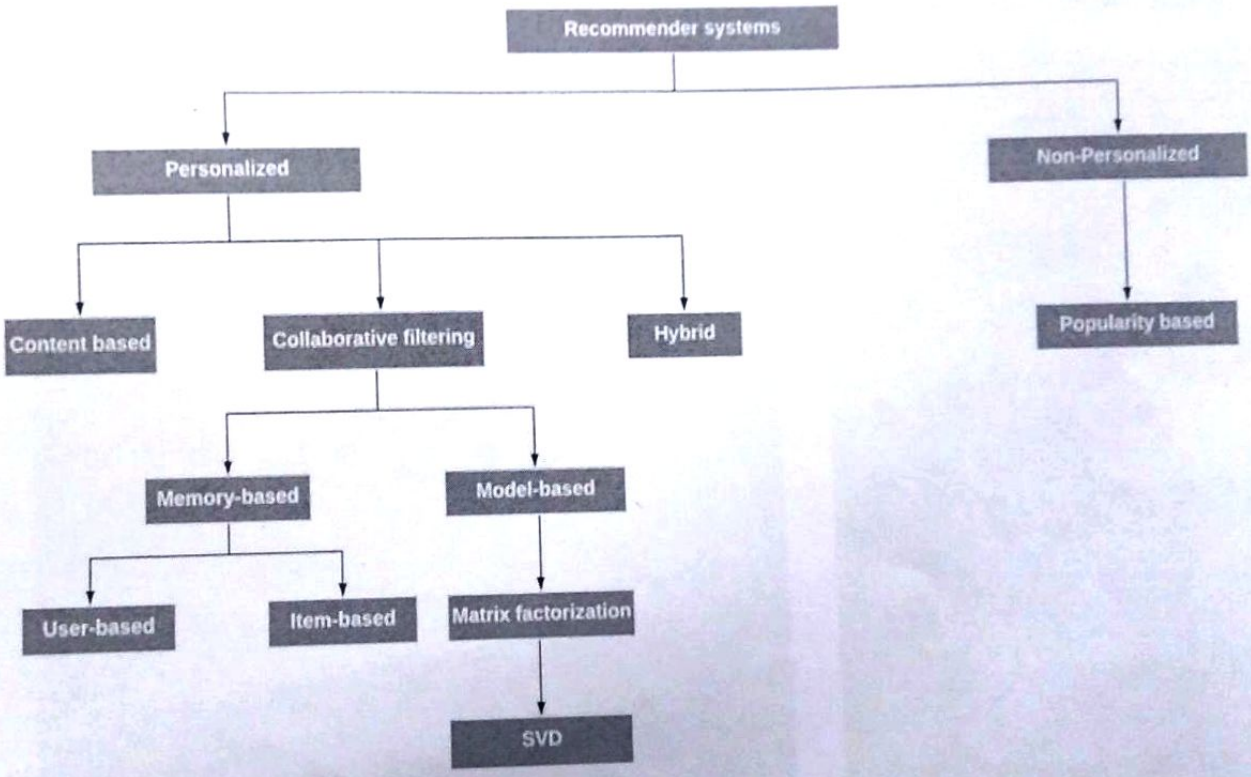


After watching Udemy online course Building Recommender Systems with Machine Learning and AI, I came up with the idea to write a text that can help beginners to understand the basic ideas of the recommender systems.

A **recommender system**, or a **recommendation system** is a subclass of information filtering system that seeks to predict the "rating" or "preference" a user would give to an item.

In the last decade companies have invested a lot of money in their development. Netflix awarded a $1 million prize to a developer team in 2009 for an algorithm that increased the accuracy of the company's recommendation engine by 10 percent.

There are two main types of recommender systems – personalized and non-personalized.



Picture 1 – Types of recommender systems

Non-personalized recommendation systems like popularity based recommenders recommend the most popular items to the users, for instance top-10 movies, top selling books, the most frequently purchased products.

What is a good recommendation?

- The one that is personalized (relevant to that user)

- The one that is diverse (includes different user interests)
- The one that doesn't recommend the same items to users for the second time
- The one that recommends available products

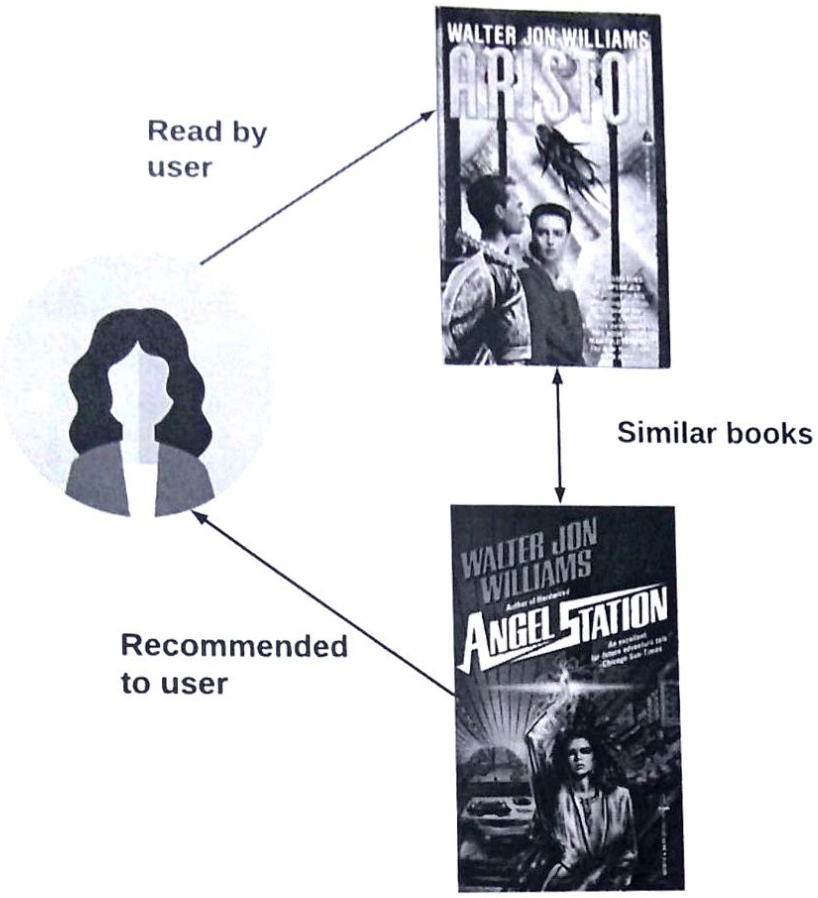
Personalized

Personalized recommender system analyzes users data, their purchases, rating and their relationships with other users in more detail. In that way every user will get customized recommendations.

The most popular types of personalized recommendation systems are content based and collaborative filtering.

Content based

Content based recommender systems use items or users metadata to create specific recommendations. The user's purchase history is observed. For example if a user has already read a book from one author or bought a product of a certain brand it is assumed that the customer has a preference for that author or that brand and there is a probability that user will buy a similar product in the future. Assume that Jenny loves sci-fi books and her favourite writer is Walter Jon Williams. If she reads the Aristoi book, then her recommended book will be Angel station, also sci-fi book written by Walter Jon Williams.



Picture 2 – Content based recommender system

Collaborative filtering in practice gives better results then content based approach. Perhaps it is because there is not as much diversity in the results as in collaborative filtering.

Disadvantages of content based approach:

- There is a so-called phenomenon filter bubble. If a user reads a book about a political ideology and books related to that ideology are recommended to him he will be in the "bubble of his previous interests".
- lot of data about user and his preferences needs to be collected to get the best recommendation
- in practice there are 20% of items that attract the attention of 70-80% of users and 70-80% of items that attract the attention of 20% of users. Recommender's goal is to introduce other products that are not available to users at first glance. In a content based approach this goal is not achieved as well as in collaborative filtering.

Collaborative filtering

The idea of collaborative filtering is simple: User group behavior is used to make recommendations to other users. Since the recommendation is based on the preferences of other users it is called collaborative.

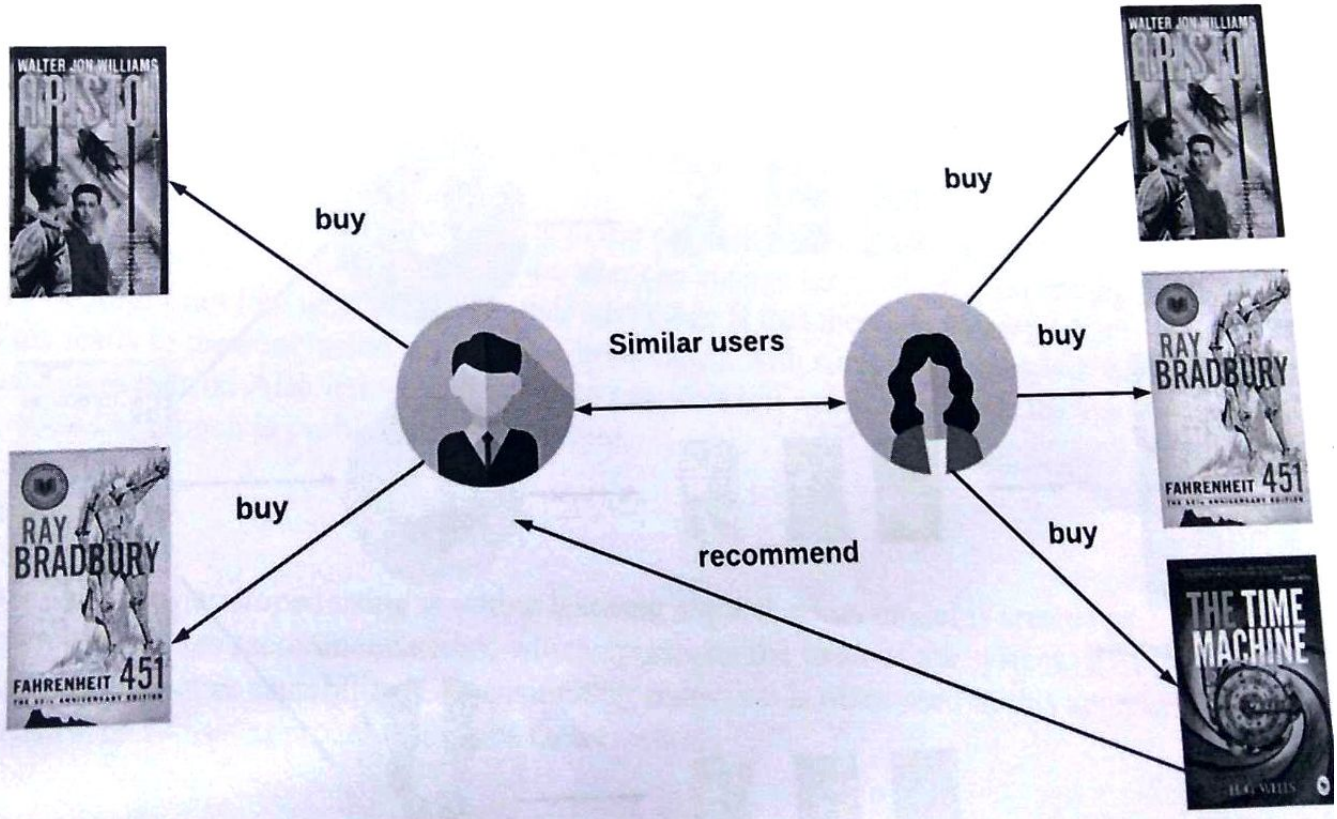
There are two types of collaborative filtering: memory-based and model based.

Memory based

Memory based techniques are applied to raw data without preprocessing. They are easy for implementation and the resulting recommendations are generally easy to explain. Each time it is necessary to make predictions over all the data which slows down the recommender.

There are two types: user based and item based collaborative filtering.

- **User based** – "Users who are similar to you also liked…" Products are recommended to the user based on the fact that they were purchased / liked by users who are similar to the observed user. If we say that users are similar what does that mean? For example, Jenny and Tom love sci-fi books. When a new sci-fi book appears and Jenny buys that book, since Tom also likes sci-fi books then we can recommend the book that Jenny bought.



Picture 3 – User based collaborative filtering recommender system