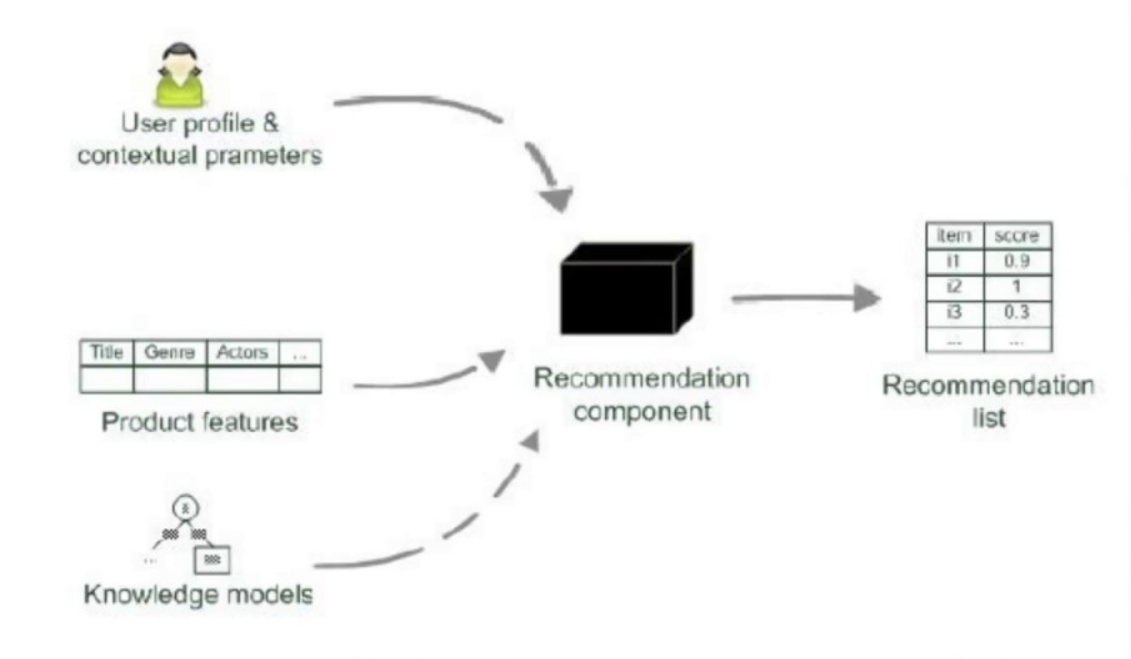
# Recommendation systems



Recommender systems are utilized in a variety of areas and are most commonly recognized as playlist generators for video and music services like Netflix, YouTube and Spotify, product recommenders for services such as Amazon, or content recommenders for social media platforms such as Facebook and Twitter.

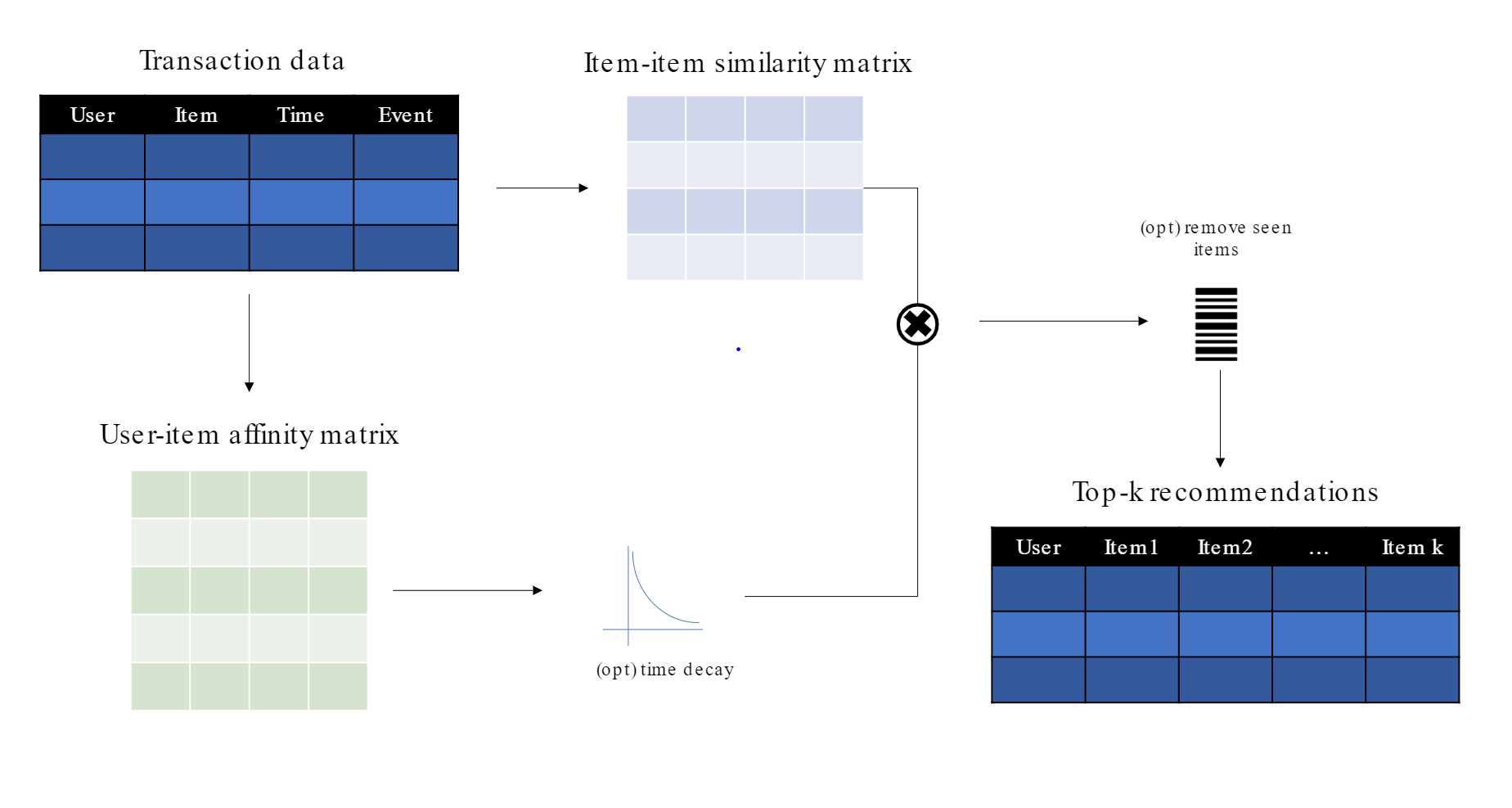
In this case- it’s for a company called SNACKFAIR that competes against online subscription services.

# The given algorithm is SAR

The SAR (Simple Algorithm for Recommendation) algorithm is a type of Collaborative Filtering algorithm which provides a fast, scalable, adaptive algorithm for personalized recommendations based on user transaction history.

It is powered by understanding the similarity between items, and recommending similar items to those a user has an existing affinity for

The algorithm provides the most relevant information to a user by discovering patterns in a dataset



# When to use the algorithm?

SAR is focussed on implicit ratings.

The SAR algorithm can be used when we have an idea about the items which the user/customer consumes. The algorithm can be used to provide recommendations of new products to buy for users based on their previous buying trend.

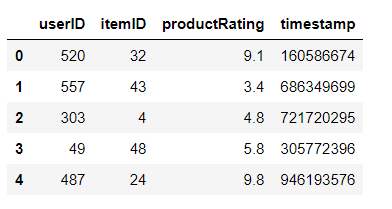
The affinity matrix in SAR captures the strength of the relationship between each individual user and the items that user has already interacted with.in the past.

So it helps in recommending the top N items for a user among recommended K items

# Sample Data

Data was generated using Faker - A python library

The schema is as shown below



The Split : Stratified split

The split is stratified so that the same set of users or items will appear in both training and testing data sets.

We also generated an addition dataset for a more detailed analysis of the product so that we can cater to the preferences of the user based on various factors



# 

# Working Prototype and Modelling

We model the algorithm with the above dataset.

The similarity type used is jaccard.

The **Jaccard similarity** compares members for two sets to see which members are shared and which are distinct. It's a measure of **similarity** for the two sets of data, with a range from 0% to 100%. The higher the percentage, the more similar the two are

SAR is a kind of neighborhood based algorithm which is intended for ranking top items for each user.

It handles cold- items and semi-cold users.

**Collaborative Filtering**

SAR is based on a simple yet powerful collaborative filtering approach. In the heart of the algorithm is computation of two matrices:

* Item-to-Item similarity matrix
* User-to-Item affinity matrix

# Insights & How SAR Algorithm helps Snackfair

Snackfair can handle cold- items and semi-cold users

* Pushing new items to active users is not a good strategy.
* Interestingly, Snackfair can connect new items with some less active users which would statistically yield better performance
* The most purchased items are the ones that are already well known to the users so advertising will not be needed for those items. Newer items are the ones to be advertised for

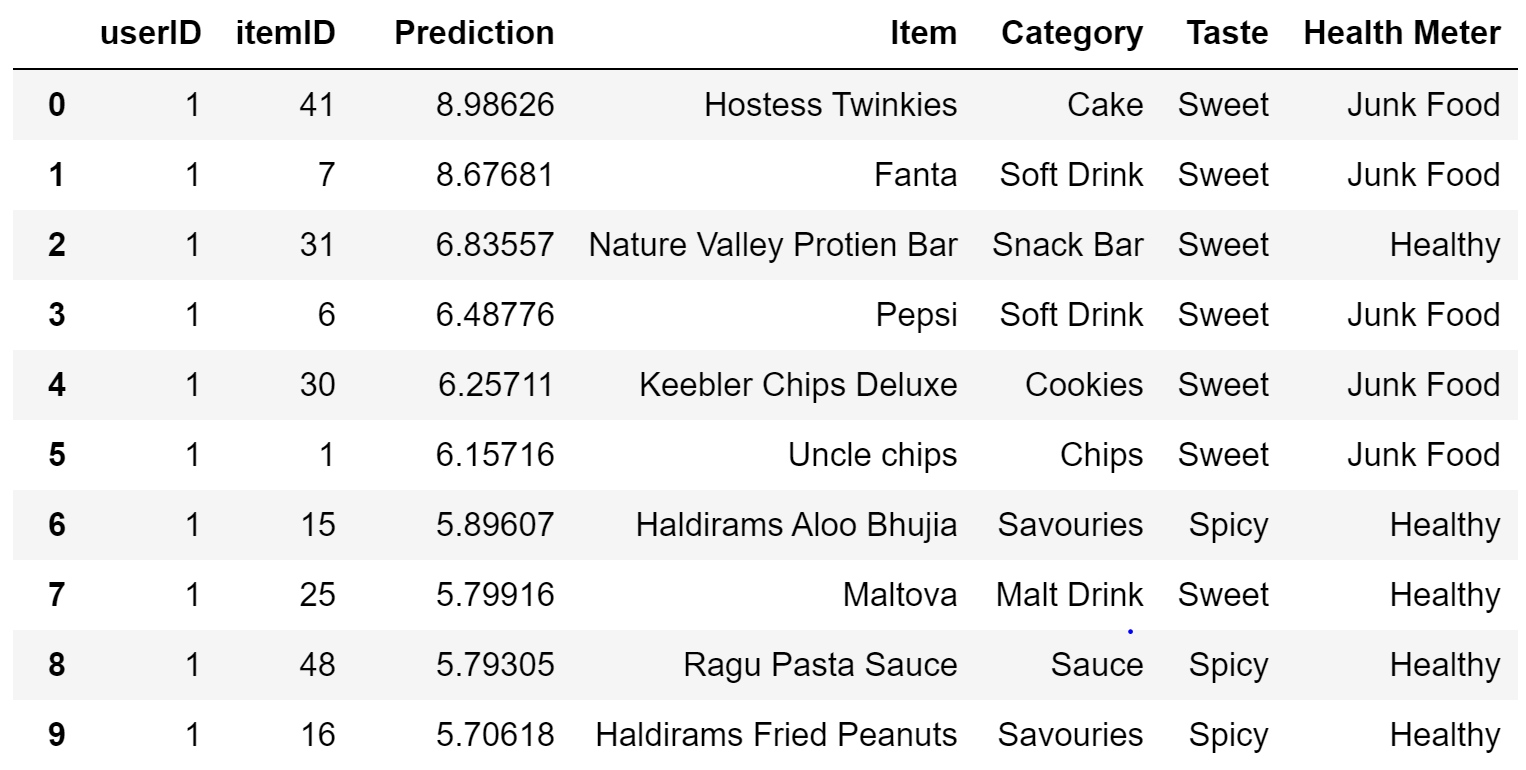
SAR favors an implicit rating scenario and it does not predict ratings

SAR Algorithm provides 2 kinds of recommendations

* User recommendations, which recommend items to individual users based on their transaction history, and
* Frequently-Occurring-Together recommendations, which recommend items similar to a given set of items.

Eg for User recommendations

Personalized recommendations for a single user are obtained by multiplying the Item-to-Item similarity matrix with a user affinity vector. Here, the jaccard similarity type method is used to calculate the item similarity



Eg for Frequently-Occurring-Together recommendations

Frequently-Occurring-Together (FOT) recommendations (also called similar-item recommendations) are based purely on the Item-to-Item similarity matrix.