

**Machine learning with pyspark**

Big data project by : SORY Anas –Favlert Ruben - FRITEL ludo

# Recommender system

Recommender systems are mainly used for auto-suggesting the right content or product to the right users in a personalized manner to enhance the overall experience.

Recommender systems are really powerful in terms of using huge amounts of data so does SPARK.  
Recommender systems can be used for multiple purposes in the sense of recommending various things to users :

* ADS ( Youtube )
* Jobs ( linked in )
* Connections/friends (linked in / facebook )
* Movies ( Netflix )

The impact of these recommendations is proving to be immense from a business standpoint. . Some of the immediate benefits that RS offer in retail settings are:

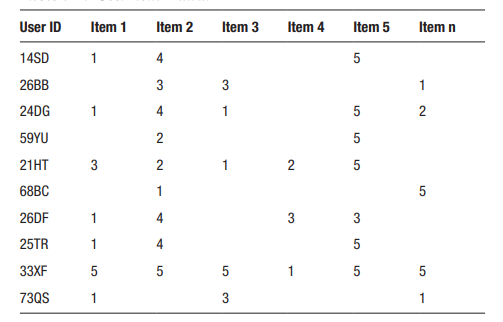
1. Increased Engagement
2. Increased Revenue
3. Positive Reviews and Ratings by Users

Now that we know the usage and features of RS, we can take a look at different types of RS. There are mainly five types of RS that can be built :

# USER ITEM MATRIX

In the rows, we have all the unique users; and along the columns, we have all the unique items. The values are filled with feedback or interaction scores to highlight the liking or disliking of the user for that product.

And example of the matrix is the following :



As we cab see, all user does not interact with all product so we have a lot of null values . The values in the matrix are generally feedback values deduced based upon the interaction of the user with that particular item.

# Missing values

The user item matrix would contain lot of missing values for the simple reason that there are lot of items and not every user interacts with each item. There are a couple of ways to deal with missing values in the UI matrix

* Replace the missing value with 0s.
* Replace the missing values with average ratings of the user.

# Hybrid recommedner systems

As the name suggests, the hybrid RS include inputs from multiple recommender systems, making it more powerful and relevant in terms of meaningful recommendations to the users.

