# Recommendation System:

A recommendation system is a subclass of Information filtering Systems that seeks to predict the rating or the preference a user might give to an item. In simple words, it is an algorithm that suggests relevant items to users.




| | SPIDERMAN | AVENGERS | COCO | MONSTERS UNIVERSITY |
|---|---|---|---|---|
| User 1 | 4 | 3 | 3 | 4 |
| User 2 | 3 | 5 | 4 | 2 |
| User 3 | 5 | 4 | 4 | 3 |
| User 4 | 2 | 4 | 3 | 3 |
| User 5 | 2 | 5 | 5 | 1 |

## Before we get started, let me define a few terms to describe the datasets:

- Item: A thing that is recommended, whether it be a product, movie, webpage, or piece of information.
- User: A person who rates items and receives recommendations for new items.
- Rating: An expression of preference by a user in regards to items. Ratings can be binary (like/dislike), integer (1 to 5 * stars), or continuous (any value on some interval). Additionally, there are implicit ratings which record only whether a user interacted with an item.

## Types of Data Used by Recommender Systems:

Since big data fuels recommendations, the input needed for model training plays a key role. Depending on your business goals, a system can work based on such types of data as content, historical data, or user data involving views, clicks, and likes. The data used for training a model to make recommendations can be split into several categories.

1) **User behavior data (historical data)**

- Log on-site activity: clicks, searches, page, and item views
- Off-site activities: tracking clicks in emails, in mobile applications, and in their push notifications

2) **Particular item details**

- Title
- Category
- Price
- Description
- Style
- Rating

3) **Contextual information**

- Device used
- Current location
- Referral URL

For you to get a full picture of your customer, it is not enough to be aware of what he or she is viewing on your website and your competitors' ones. You should take into account the frequency of visits, user location, and types of devices. All the data sources are equally important for the smooth and consistent operation of different types of algorithms.

Owning this information brings you closer to a 29% increase in sales. That is precisely what Amazon experienced firsthand after they had implemented recommendation engines on their website.

But if you want to take content or user features into account, you need to deal with various types of data. That will require a best-fit algorithm, and you will have to solve data-specific tasks. Besides, in case you are launching a new service not having historical data – in other words, in case of cold start – content analysis is everything you have. To make things less complicated, contact us, and our experts into ML and big data will help you.

## Which datasets i will use in this project?

I am going to use two datasets for recommender system. Movie dataset and amazon review dataset. These dataset does not contains much features and are very clean so no need for detailed preprocessing.

**Movies rating dataset features:** 1) UserId
2) MovieId
3) Rating (Target column)
4) Title

**Products rating dataset features:**

1) UserId
2) ProductId
3) Rating (Target column)
4) Timestamp