**Problem Statement -**

Build your own recommendation system for products on an e-commerce website like Amazon.com.

Online E-commerce websites like Amazon, Filpkart uses different recommendation models to provide different suggestions to different users.

**Amazon** currently uses item-to-item collaborative filtering, which scales to massive data sets and produces high-quality **recommendations** in real time. This type of filtering matches each of the user's purchased and rated items to similar items, then combines those similar items into a **recommendation** list for the user.

In this project we are going to build recommendation model for the electronics products of Amazon.

The dataset here is taken from the below website.

**Source** - Amazon Reviews data (http://jmcauley.ucsd.edu/data/amazon/) The repository has several datasets. For this case study, we are using the Electronics dataset.

**Dataset columns** - first three columns are userId, productId, and ratings and the fourth column is timestamp. You can discard the timestamp column as in this case you may not need to use it.

Please do the analysis based on steps( 1 to 8) as given below -

**Steps -**

1. Read and explore the given dataset. ( Rename column/add headers, plot histograms, find data characteristics)
2. Take a subset of the dataset to make it less sparse/ denser. ( For example, keep the users only who has given 50 or more number of ratings )
3. Split the data randomly into train and test dataset. ( For example, split it in 70/30 ratio)
4. Build Popularity Recommender model.
5. Build Collaborative Filtering model.
6. Evaluate both the models. ( Once the model is trained on the training data, it can be used to compute the error (RMSE) on predictions made on the test data.)
7. Get top - K ( K = 5) recommendations. Since our goal is to recommend new products to each user based on his/her habits, we will recommend 5 new products.
8. Summarise your insights.

**Hints -**

Dataset Link - https://drive.google.com/file/d/1ClBptsK3V5KgKXtK2GSRzFNAW7GnTPDW/view?usp=sharing

Please Note -

● If you are facing any memory issue while working on this project, create a small subset of data and work on it.

● If you are stuck at model evaluation part of this project.

Please refer to below links -

1.<https://surprise.readthedocs.io/en/stable/accuracy.html>

2.<http://surpriselib.com/> - Getting started, example

