**Collaborative Filtering Methods In Product Recommendation**

Nowadays, Product Recommendation is a widely used tool in a variety of areas from music apps to e-shops or pretty much in any platform the user interacts with. Recommendation Systems’ purpose is to attract new customers or to help the current ones by offering suggestions based on habits, interests, prior transactions or even the customer's profile.

Recommendation Systems are classified based on:

* What data the recommender engine needs as to make the recommendations?
* How the recommendations are produced?

There are two main categories, **content** filtering and **collaborative** based filtering.

In **content** based filtering, the data used, include customer as well as product metadata information. According to the genre eg.in movies: If a customer watched a romantic movie in the past, the system is going to assume that the customers "likes" romance genre, so the output will include movies that have been categorized with the romance label.

Another way, to make recommendations, is by looking into the customer's transactions (interactions with available products) and try to predict what "rating" the customer would give to an unrated product. **Collaborative** filtering does that, not by looking into categories and tags, but by trying to find customers that have interacted with the same products our target customer has (user based recommendations). These similar customers and their preferences are the source for the recommender engine as to predict for every product the target customer has not interacted with, what rating would he give. For example, in gaming: Paul bought FIFA 2021 and Jo also bought FIFA 2021, Jo also bought Super Mario, so the systems recommends Super Mario to Paul.

Another approach, for making this prediction comes from finding products displaying a degree of similarity to the products the target customer has already purchased (item based recommendations), eg.in groceries: It is quite common for someone who bought milk to buy cereal as well. The difference is that now we are focusing on products rather than customers to help us make our recommendations.

This project looks into **Collaborative Filtering Methods In Product Recommendation.**

The goals were to:

* Understand how a recommendations systems works.
* Which method suit the data we are provided with? (collaborative, content, hybrid)
* How important is the "rating" as to recognize the customer, product interactions.
* Which algorithms and methods produce the best output. (Surprise, LightFM library)
* How to evaluate a product recommendation system?
* Are the recommendations produced by this system aligned with the customer's needs?

For the technical part, as to deliver answers to the above goals two different Python libraries were used:

* Surprise
* LightFM

**Project Structure:**

* Prior to any model building, looking into our data in the **Data\_Preprosessing.ipynb** Jupyter Notebook was a must, as to understand and adjust the data according to the project’s needs.
* Then, in **RS1\_Surprise.ipynb** and by using the Surprise library models it provides, recommendations were produced, evaluated and compared across the different models and model hyperparameter tunings.
* The same process took place for LightFM and its embeddings model in **RS2\_LightFM.ipynb** .
* In **Comparison.ipynb**, both Surprise and LightFM results are compared as to find out which system outperformed the other.
* Also, All\_Functions.py is a python file, with all the libraries, packages used and the customed made functioned created for this project.
* It does not run separately, but as part of the delivered Jupyter Notebooks, for functionality reasons.

**To sum up, this assignment focuses on Collaborative Filtering. The goal is to produce User Based Recommendations. The data available are transactions made by foodservice customers (Everest) dated from 2020-01-01 to 2020-07-31.**