

## NutriGro (Building a Recommender System)

This project aims to develop a recommender system that provides personalized product recommendations. The system leverages historical order data and healthiness information to suggest products that align with user preferences and health goals. A content-based recommender system provides users with suggestions based on similarity in content. This series will demonstrate a popular solution to this problem, recommender engines! What is a recommender engine? you may ask. It is an unsupervised learning algorithm (one that does not have a target variable to measure accuracy against) mostly used to aid in consumer decision making. I'm sure you have seen them while online shopping. They also appear in places like streaming apps (aka Netflix and Hulu) to help you select a TV show or movie to watch next and on journalism/media websites like Medium to suggest other articles you may like to read, among many other uses.

### Data Sources

The dataset for this project comprises data interaction and data recipe.

### Data Preprocessing

Once the data is collected, the next critical step is data pre-processing. This involves cleaning the data to remove inaccuracies or missing values and transforming it into a structured format, such as CSV or JSON, that can be easily processed. A critical phase in which the collected data is refined and prepared for analysis. First, data cleansing is performed to remove irrelevant, incomplete, or erroneous information. This may involve filtering out noise or correcting data inconsistencies to ensure that the remaining data is accurate and reliable. Clean and preprocess the historical order and healthiness data. Handle missing values and encode categorical variables. Data normalization is the process of transforming data to a common scale, which helps to compare and analyze data that was originally in different formats or scales.

### Filtering

Matrix factorization is a mathematical technique for predicting user preferences. It works by breaking down a large user-item interaction matrix into smaller, more manageable matrices representing users and items. These matrices are then used to identify latent factors that influence user preferences. By applying specific mathematical recommendation algorithms, the system can predict how likely a user is to prefer an item, even if they haven't interacted with it before.

### Generating

A crucial phase in which the processed data and the insights gained from the previous steps are used to suggest relevant items or content to the user.

In this stage, the engine applies algorithms to predict and match user preferences with available items to provide personalized and relevant suggestions

The engine considers factors such as past user behavior, similarities between items, and user profiles to generate these recommendations.

**Personalized Recommendations:** A recommender tool tailored specifically to an individual's preferences and past behavior, these suggestions are based on items the user has previously interacted with, showing similar or complementary products.

**Best Sellers:** These are popular items across the platform, often recommended to new users or those with limited interaction history. They represent what is trending or most purchased in a certain category.

**Related Items:** Often seen as "Customers who viewed this also viewed" suggestions, these are based on the correlation between products, recommending items that other users have looked at or purchased in relation to the current item.

**New Arrivals:** Recommendations focusing on the newest items in a category, useful for returning users to discover the latest products or content.

## Improved Customer Experience

Personalized recommendations enhance the user experience by making it easier for customers to find products or content that match their interests.

This personal touch makes the browsing experience more engaging and satisfying, leading to increased customer satisfaction and loyalty.