Recommender System

Personalized Item Recommender that studies your history to predict your preferences

Introduction

How many decisions do you make in a day? An average person believes there are at least a few dozens of conscious decisions made every day. Some sources say people make around 35000 decisions each day. Like sleeping, eating or even breathing, our decision making abilities are being used whether we like it or not. More precisely put, decisions are made whether we've had enough time to make the right decision or not.

This Recommender System hopes to assist in making purchase decisions and to save research time.

The problem

Out of roughly 35000 decisions made in a day, let's keep personal decisions personal. Deciding on what to say to your boss tonight is not what is at stake here.   Deciding how to raise your child is your decision to make.

We are asking the questions like, what are we eating for lunch tomorrow? Questions like, which movie should I watch next? Which type of sound speaker should a rock band guitarist purchase. And if the buyer was a construction worker?

It is very easy to go to a bookstore 'horror' section and pick up the book next to the reader's favorite horror book. But the only connection between the first book and the second is that they are both classified as 'horror' and there is only a slim chance a reader

will like all horror books . The probability that a reader will like the second horror book becomes pretty low.

So what are the predictors to keep track of? What sort of information is critical when an individual makes a purchase decision? Our goal is to gather these data and come up with a personalized list of items that a certain individual may have interests in.

This is for you

This analysis report will provide most insights to any ecommerce companies where there are sales and purchases. Data such as purchase history, user rating, user info, item info are valuable for a problem like this and ecommerce companies are built on these structures making them very suitable to take advantage of this project.   This may also help entertainment platform companies that have copyrights to intellectual properties looking to research market and user preference demographics. This system will also be able to incorporate into the platform to produce recommendations for the users. Lastly, this is for any research for consumer behaviors and individuals looking for recommendations based on one's preferences.

The Data

Data used in this project will come from many different sources. One of the goals for this project is to have the recommender be adaptable to different items and industries. So this recommender will require certain structure of input data and those data that are generic and available to most items and industries.

http://deepyeti.ucsd.edu/jianmo/amazon/index.html

Recommender Systems Dataset by Julian McAuley, UCSD

Contains numerous links to product and user review data like Amazon, books, fashion and food. Amazon data will be primarily used.

http://cseweb.ucsd.edu/~jmcauley/datasets.html

Public Dataset for Recommender Systems

Contains numerous links to ecommerce data like Amazon, Apple music, and movie data.

Big company data will be primarily used for details regarding products and the number.

https://github.com/caserec/Datasets-for-Recommender-Systems

Trajectory

Initially, there will be extensive data transformation to turn raw data into usable format. Our goal is to create a generic structure of input data and a model that is versatile enough to be used in different industries. So our data will most likely consist of defining features like item, item descriptions, item rating, number of ratings, item category and user data like purchase of history, rating for the purchases, user description and etc.   These forms of data are commonly used in ecommerce and entertainment platform so it will be optimal.

Now, there are different forms of developing a recommender system.

- Demographic
    - This is the most basic recommender where model looks at the popularity score and suggests the most popular item.

        This does not capture the personalized preferences.

- Content-based Filtering
    - This method looks at the item information of prior purchase and recommends a similar item.

        This is a limited personalization as the model only guess the similar products

- Collaborative Filtering
    - This method looks at other user's purchase history and recommends the similarly liked item

All these different types of methods show strengths and weakness. The Recommender System developed in this project will be a hybrid of all of the above methods to deliver the best accuracy.

This trajectory is subject to change as the project goes on.

The Deliverables

By the end of this project, the deliverables will include the report itself, original dataset, transformed dataset, code used, and slide deck presentation of the findings.

Links used for research:

https://towardsdatascience.com/how-to-build-a-simple-recommender-system-in-python-375093c3fb7d

https://towardsdatascience.com/make-your-own-recommendation-system-b596d847296d

https://awesomeopensource.com/projects/recommender-system

https://www.kaggle.com/ibtesama/getting-started-with-a-movie-recommendation-system#The-Age-of-Recommender-Systems