SOCIAL NETWORK ANALYSIS PROJECT

Product Recommendation System for e-commerce

Abstract

In today's world, we find a wide variety of search options and we may have difficulty selecting what we really need. The recommendation System plays an important part in dealing with these problems. A recommender system is a framework that is a filtering system that filters the data with various algorithms and recommends the user with the most relevant data. Recommendation systems are productive customization mechanisms, often up-to-date and recommendations based on current consumer preferences. These systems have shown to be extremely helpful in different areas of e-commerce, education, movies, music, books, films, scientific papers, and various products. In this project Product Recommendation system for e-commerce is implemented. Product-recommendation system is powered by machine learning and employs suggestions of goods associated with a brand's digital assets. Driven by a range of algorithmic decisions, recommendations algorithms have a customized experience for user, product and background information. This helps people find what they want and the items they're searching for thus improving the search process. In this way, enterprises can understand more about the individual needs and desires of a customer optimize success in real time while improving their long-term research roadmaps

About Dataset

Dataset name: Amazon Product review dataset

This is a dataset related to over 2 Million customer reviews and ratings of products sold on their website.

It contains-
- the unique UserId (Customer Identification),
- the product ASIN (Amazon's unique product identification code for each product),
- Ratings (ranging from 1-5 based on customer satisfaction) and
- the Timestamp of the rating (in UNIX time)

1	UserId	ProductId	Rating	Timestamp
2	A39HTATAQ9V7YF	205616461	5	1369699200
3	A3JM6GV9MNOF9X	558925278	3	1355443200
4	A1Z513UWSAAO0F	558925278	5	1404691200
5	A1WMRR494NWEWV	733001998	4	1382572800
6	A3IAAVS479H7M7	737104473	1	1274227200
7	AKJHHD5VEH7VG	762451459	5	1404518400
8	A1BG8QW55XHN6U	1304139212	5	1371945600
9	A22VW0P4VZHDE3	1304139220	5	1373068800
10	AVOGV98AYOFG2	1304168522	5	1372118400
11	A6R426V4J7AOM	1304168522	5	1373414400
12	A22VW0P4VZHDE3	1304174778	5	1372896000
13	AKGB62WGF35J8	1304174778	5	1372896000
14	A22VW0P4VZHDE3	1304174867	5	1373068800
15	A1BG8QW55XHN6U	1304174867	5	1372291200
16	A1BG8QW55XHN6U	1304174905	5	1372291200
17	A22VW0P4VZHDE3	1304196046	5	1372896000
18	A22VW0P4VZHDE3	1304196062	5	1372896000
19	A3A4C2K3TWDAO5	1304196070	1	1378425600
20	A3FV2Q7WPZMQPV	1304196135	5	1375488000
21	A22VW0P4VZHDE3	1304196135	5	1372896000

Dataset name: Home Depot Product Search Relevance

File descriptions
train.csv - the training set, contains products, searches, and relevance scores
test.csv - the test set, contains products and searches. You must predict the relevance for these pairs.
product_descriptions.csv - contains a text description of each product. This table can be joined to the training or test set via the product_uid.
attributes.csv - provides extended information about a subset of the products (typically representing detailed technical specifications). Not every product will have attributes.
sample_submission.csv - a file showing the correct submission format
relevance_instructions.docx - the instructions provided to human raters

- The relevance is a number between 1 (not relevant) to 3 (highly relevant).
- For example, a search for "AA battery" would be considered highly relevant to a pack of size AA batteries (relevance = 3), mildly relevant to a cordless drill battery (relevance = 2), and not relevant to a snow shovel (relevance = 1).

Data fields

- **id** - a unique Id field which represents a (search_term, product_uid) pair
- **product_uid** - an id for the products
- **product_title** - the product title
- **product_description** - the text description of the product (may contain HTML content)
- **search_term** - the search query
- **relevance** - the average of the relevance ratings for a given id
- **name** - an attribute name
- **value** - the attribute's value

Our contribution

The recommendation system is designed in 3 parts based on the business context:

- **Recommendation system part I:** Product popularity based system targeted at new customers
- **Recommendation system part II:** Model-based collaborative filtering system based on customer's purchase history and ratings provided by other users who bought items similar items
- **Recommendation system part III:** When a business is setting up its e-commerce website for the first time without any product rating

Thank you