

PRODUCT RECOMMENDATION SYSTEM WITH FLASK API

FATMA BETÜL AKYOL

Email: akyolfbetul@gmail.com

1. PURPOSE AND SCOPE

The aim of this project is to make a recommendation system that is about recommending similar products to the user that has an item in the cart. The products that are recommended by the system should be similar to item in the cart. Ten products will be recommended to the user.

2. MATERIALS and METHODS

The project was developed using Python programming language. The project is split into two subcategories:

1. *Product recommendation*
2. *API generation*

2.1 Product Recommendations

In the project recommendation step two datasets were explored. These datasets are:

Events.json: Contains cart operation with product id and price information.

Meta.json: It includes detailed information about products.

First five rows of the events and meta files are shown in the Figure 1.

	event	sessionid	eventtime	price	productid
0	cart	a0655eee-1267-4820-af21-ad8ac068ff7a	2020-06-01T08:59:16.406Z	14.48	HBV00000NVZE8
1	cart	d2ea7bd3-9235-4a9f-a9ea-d7f296e71318	2020-06-01T08:59:46.580Z	49.90	HBV00000U2B18
2	cart	5e594788-78a0-44dd-8e66-37022d48f691	2020-06-01T08:59:33.308Z	1.99	OFIS3101-080
3	cart	fdfeb652-22fa-4153-b9b5-4dfa0dcaffdf	2020-06-01T08:59:31.911Z	2.25	HBV00000NVZBW
4	cart	9e9d4f7e-898c-40fb-aae9-256c40779933	2020-06-01T08:59:33.888Z	9.95	HBV00000NE0T4

(a)

	productid	brand	category	subcategory	name
0	HBV00000AX6LR	Palotte	Kişisel Bakım	Saç Bakımı	Palotte Kalıcı Doğal Renkler 10-4 PAPATYA
1	HBV00000BSAQG	Best	Pet Shop	Kedi	Best Pet Jöle İçinde Parça Etli Somonlu Konser...
2	HBV00000JUHBA	Tarım Kredi	Temel Gıda	Bakliyat, Piring, Makarna	Türkiye Tarım Kredi Koop.Yeşil Mercimek 1 kg
3	HBV00000NE0QI	Namet	Et, Balık, Şarküteri	Şarküteri	Namet Fıstıklı Macar Salam 100 gr
4	HBV00000NE0UQ	Muratbey	Kahvaltılık ve Süt	Peynir	Muratbey Burgu Peyniri 250 gr

(b)

Figure 1 (a) Events.json (b) meta.json

As a next step, the tables which are events and meta were joined. The inner join was used as the join method and the join operation is done based on the productid feature placed in the two tables.

To find items that are taken in one session is important for the recommendation case. For this reason, a pivot table that contains sessions and products were created. Index values of the table are session ids and columns are product ids.

```
product_pivot = pd.pivot_table(joined_df, index = 'sessionid', columns = 'productid', values = 'category', aggfunc = 'count')
```

In order to search items that were bought together a co-occurrence matrix was created. If items were bought together the value of these items was assigned to 1. Calculation of the co-occurrence matrix is done with the dot product of the pivot table and the transpose of the pivot table.

```
co_occurrence = product_pivot.T.dot(product_pivot)
np.fill_diagonal(co_occurrence.values, 0)
```

The diagonal of the matrix is the intersection of the same item in the index and the column. So, to prevent ambiguity in the similarity calculation these values were filled with 0. The first five rows of the co-occurrence matrix are shown in Figure 2.

	productid	AILEBIZISMTLDGY54	AILEBIZISMTLDHB18	AILEBSHSB22037	AILEDALIN275107	AILEELITDIS5285B	AILEELITDIS5288B	AILEFIMAN5299094	AILEHIPPTR2477
productid									
AILEBIZISMTLDGY54		0.0	1.0	1.0	0.0	1.0	1.0	0.0	0.0
AILEBIZISMTLDHB18		1.0	0.0	1.0	0.0	1.0	1.0	0.0	0.0
AILEBSHSB22037		1.0	1.0	0.0	0.0	1.0	1.0	0.0	0.0
AILEDALIN275107		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AILEELITDIS5285B		1.0	1.0	1.0	0.0	0.0	1.0	0.0	0.0

Figure 2. Co-occurrence matrix

According to table in the Figure 2 products that ‘AILEBIZISMTLDGY54’ and ‘AILEBIZISMTLDHB18’ were bought together. 0.0 values refer to the items that were not taken in the same session.

2.1.1 Similarity Calculation

The co-occurrence matrix was created for finding bought-together items. In order to find similarities between items, we need a calculation of similarity. Cosine similarity was chosen as a similarity metric.

```
from sklearn.metrics.pairwise import cosine_similarity
similarity_df = pd.DataFrame(cosine_similarity(co_occurence))
```

As a result of similarity calculation is shown in the Figure 3.

	AILEBIZISMTLDGY54	AILEBIZISMTLDHB18	AILEBSHSB22037	AILEDALIN275107	AILEELITDIS5285B	AILEELITDIS5288B	AILEFIMAN5299094	AILEHIPPTR2477
productid								
AILEBIZISMTLDGY54	1.000000	0.991667	0.991667	0.0	0.991667	0.991667	0.037268	0.0
AILEBIZISMTLDHB18	0.991667	1.000000	0.991667	0.0	0.991667	0.991667	0.037268	0.0
AILEBSHSB22037	0.991667	0.991667	1.000000	0.0	0.991667	0.991667	0.037268	0.0
AILEDALIN275107	0.000000	0.000000	0.000000	1.0	0.000000	0.000000	0.000000	0.0
AILEELITDIS5285B	0.991667	0.991667	0.991667	0.0	1.000000	0.991667	0.037268	0.0

Figure 3. Cosine similarity results

Ten similar products will be recommended to the user. So, the similarity values were sorted in descending order.

```
ten_products.append(similarity_df[similarity_df.index!=i][i].sort_values(ascending = False)[0:10].index)
ten_products_df = pd.DataFrame(ten_products)
```

In order to ignore similarity values for the same item, a condition was added to the code snippet above.

2.2. API Generation

Product recommendation system was served with API. Flask was used for API generation. Flask codes were written separate python file in the project directory. To get similar products *find_product* method that located in recommendation code was called. A specific product ID was sent to *find_product* method. Return values were sent in the json format to the API. The application is run on port number 5004 and the route products page. URL address of the application is <http://localhost:5004/products>.

```
@app.route('/products', methods = ['GET'])
def get_product():
    print("product_id: " + str('OFISDUR21'))

    return hburada.find_products('OFISDUR21')

if __name__ == '__main__':
    app.run(host='localhost', port=5004)
```

Figure 4. Flask API

Examples of return values of API is shown Figure 5 (a), (b) and (c).

```
{
  "name": "Milka Üzüm Fındıklı Tablet Çikolata 80 gr",
  "price": "7.95",
  "name": "Godiva Lokma Fındıklı Çikolata 130 g",
  "price": "23.00",
  "name": "ZYULUDAMSD009",
  "name": "Uludağ 250 Ml Frutti Extra Mandalina",
  "price": "2.95",
  "name": "ZYHPSOLENKG006",
  "name": "Ozmo Cornet 25 gr",
  "price": "1.50",
  "name": "HBV00000TWJKA",
  "name": "Fondante Çikolata Dolgulu Naneli Fondan Kutu 300 g",
  "price": "10.90",
  "name": "HBV00000PQJLZ",
  "name": "Eti Crax Cubuk Kraker 52 gr",
  "price": "0.50",
  "name": "HBV00000PQJLX",
  "name": "Eti Çikolata Uzun Antep Fıstıklı 34 Gr",
  "price": "2.10",
  "name": "HBV00000PVBVC",
  "name": "Rocco Sıka Şeker Ahududu-Limon 12 Gr",
  "price": "1.75",
  "name": "HBV00000COROC",
  "name": "Uludağ Frutti Extra Limon 250 Ml",
  "name": "Uludağ Frutti Extra Limon 250 Ml",
  "price": "2.95",
  "price": "2.95"
}
```

(a)

```
{
  "name": "Dentastix Small 100 g 3'lü",
  "price": "7.25",
  "name": "Enjoy Köpekler İçin Dana Etli Ödül Çubuğu 11 gr 3'Lü",
  "price": "9.75",
  "name": "HBV00000PQKQ8",
  "name": "Nestle Nesfit Gevrek Granola Kırmızı Meyveli 300 Gr",
  "name": "Nestle Nesfit Gevrek Granola Kırmızı Meyveli 300 Gr",
  "price": "16.95",
  "price": "16.95",
  "name": "HBV00000SF26G",
  "name": "Sek Protein Süt Kakaolu 330 ml",
  "name": "Sek Protein Süt Kakaolu 330 ml",
  "price": "5.95",
  "price": "5.95",
  "name": "HBV00000LMMY4",
  "name": "Pınar Kefir 1000ml",
  "name": "Pınar Kefir 1000ml",
  "price": "7.95",
  "price": "7.95",
  "name": "HBV00000EIRGB",
  "name": "Pantene 3 Minute Miracle Saç Bakım Kremi Superfood Güç ve Güçlü 200 ml",
  "name": "Pantene 3 Minute Miracle Saç Bakım Kremi Superfood Güç ve Güçlü 200 ml",
  "name": "Pantene 3 Minute Miracle Saç Bakım Kremi Superfood Güç ve Güçlü 200 ml",
  "price": "22.90",
  "price": "22.90",
  "price": "16.03",
  "price": "22.90",
  "name": "HBV00000PLGSI",
  "name": "Sek Quark Yaban Mersini 140 G",
  "name": "Sek Quark Yaban Mersini 140 G",
  "price": "4.95",
  "price": "4.95",
  "price": "4.95",
  "price": "4.95",
  "price": "4.95",
  "price": "4.95",
  "name": "HBV00000SP87P",
  "name": "Polonez Hindi Fıme 2 x 120 gr",
  "name": "Polonez Hindi Fıme 2 x 120 gr",
  "name": "Polonez Hindi Fıme 2 x 120 gr",
  "price": "25.90",
  "price": "9.90",
  "price": "25.90",
  "name": "HBV00000I6DT7",
  "name": "Elidor Doğanın Enerjisi Hindistan Cevizi Tüp Saç Kremi 200 ml",
  "name": "Elidor Doğanın Enerjisi Hindistan Cevizi Tüp Saç Kremi 200 ml",
  "price": "24.90",
  "price": "24.90"
}
```

(b)

```
{
  "name": "Pantene Saç Bakım Kremi Superfood Güç ve Güçlü 470 ml",
  "price": "17.90",
  "name": "Unabella Mini Cepli Ekmek 6'lı",
  "price": "10.50",
  "name": "HBV00000NGSRB",
  "name": "Carrefour Elma Kokulu Bulaşık Deterjanı 750 ml",
  "name": "Carrefour Elma Kokulu Bulaşık Deterjanı 750 ml",
  "price": "4.75",
  "price": "4.75",
  "name": "HBV00000IYDT",
  "name": "Dolcia Muzlu Puding 4x110 g",
  "name": "Dolcia Muzlu Puding 4x110 g",
  "price": "6.95",
  "price": "6.95",
  "name": "HBV00000BT5B7",
  "name": "Pepsi 4x1 L",
  "name": "Pepsi 4x1 L",
  "price": "16.25",
  "price": "16.25",
  "price": "16.25",
  "price": "16.25",
  "name": "HBV00000PGLQ",
  "name": "Milka Bubbly Sütü 80 g",
  "name": "Milka Bubbly Sütü 80 g",
  "name": "Milka Bubbly Sütü 80 g",
  "name": "Milka Bubbly Sütü 80 g",
  "price": "7.95",
  "price": "7.95",
  "price": "7.95",
  "price": "7.95",
  "name": "HBV00000U2B0Z",
  "name": "Protex Sıvı Sabun 700 ml U İtra Koruma",
  "name": "Protex Sıvı Sabun 700 ml U İtra Koruma",
  "price": "19.90",
  "price": "19.90",
  "name": "ZYUNIL70021126",
  "name": "Knorr Çabuk Çorba Kremalı Tavuk 18 gr",
  "name": "Knorr Çabuk Çorba Kremalı Tavuk 18 gr",
  "name": "Knorr Çabuk Çorba Kremalı Tavuk 18 gr",
  "name": "Knorr Çabuk Çorba Kremalı Tavuk 18 gr",
  "name": "Knorr Çabuk Çorba Kremalı Tavuk 18 gr",
  "price": "1.75",
  "price": "1.75",
  "price": "1.75",
  "price": "1.75",
  "name": "HBV00000IZ3MJ",
  "name": "Derby Lady 2 Bıçaklı 5'li Poşet",
  "name": "Derby Lady 2 Bıçaklı 5'li Poşet",
  "price": "6.95",
  "price": "6.95"
}
```

(c)

Figure 5. Return 10 products for (a) ZYBICN9286520- Sek Ayran (b) AILEDALIN275107- Dalin Şampuan Klasik 750+250 ML (c) SGPAN971540- Pantene Aqualight 500 ml Şampuan

3. RESULTS and DISCUSSION

In this project, a product recommendation system has been made with API. There were approximately 330.000 events in the events.json file. To use all data was very hard for me. In Colab platform session is crashed because of data size. For this reason, 50.000 data was used as event data. All of the product data was used. Another cons for me, Flask was not worked in the Jupyter notebook. As a solution to this, I prefer the Spyder notebook. The pros of the project are data is suitable for processing and usage for the recommendation. Also, the data was clear for understanding.

In the recommendation step, bought-together items were found with a co-occurrence matrix and similarity was calculated for each item pairs with cosine similarity. The similarity values were sorted in descending order.

After the recommendation step, the API step was developed with Flask. Ten similar products were shown for the given product ID is listed in the JSON format on <http://localhost:5004/products> .