E-COMMERCE RECOMMENDATION SYSTEM

A recommendation system models with multiple function on Amazon Dataset

Fiona, Isaac, Nicole, Todd

TABLE OF CONTENTS

01

BUSINESS OVERVIEW

We are

Consultant that target on local SME in providing *research and recommendation system solution*

By providing a simplistic recommendation system in a lower cost, help SME to utilize customers data they have in hand and boost their sales revenues.

Business Values

PROBLEM

- Need insights on shopping or what to sell? Or want give out recommendation to your customer?

SOLUTION

- You need a system that provide recommendation and related reviews for your references

Working Flow

40%

Data Processing

- Pandas
- Regex

50%

Sentiment Analysis

- Vadar

80%

Recommender System

- TFIDF
- Cosine Similarity

100%

Web Deployment

- Flask
- HTML
- Bootstrap

Data Preprocessing

1) **Handle Nan values**
   - remove columns in where too many Nan values and we don't need

2) **Transform a single "price" column to four separated columns and**

   make a "price_tag" column for different price ranges for easier handling

	na_num	na_percent	num_of_unique_values
customer_questions_and_answers	9086	90.86	910
sellers	3082	30.82	6581
items_customers_buy_after_viewing_this_item	3065	30.65	6749
number_available_in_stock	2500	25.00	89
price	1435	14.35	2625
customers_who_bought_this_item_also_bought	1062	10.62	8755
number_of_answered_questions	765	7.65	19
amazon_category_and_sub_category	690	6.90	255
description	651	6.51	8514
product_description	651	6.51	8514
product_information	58	0.58	9939
customer_reviews	21	0.21	9901
number_of_reviews	18	0.18	194
average_review_rating	18	0.18	19
manufacturer	7	0.07	2651
uniq_id	0	0.00	10000
product_name	0	0.00	9964

price
£269.00 - £699.99
£6.99 - £10.49
£22.39 - £41.23

lower_price	upper_price	avg_price	price_tag
22.95	22.95	22.95	below_or_equal_twentyfive
39.99	39.99	39.99	twentyfive_to_fifty
43.99	43.99	43.99	twentyfive_to_fifty
49.81	49.81	49.81	twentyfive_to_fifty
21.20	21.20	21.20	below_or_equal_twentyfive

Data Preprocessing

Hobbies > Model Trains & Railway Sets > Rail Vehicles > Trains

Hobbies , Model Trains & Railway Sets , Rail Vehicles , Trains

3) **Separate the one-line "category" column into several sub-categories tokens**

4) **Only extract the seller info from the each product for further analysis**

{"seller"=>[{"Seller_name_1"=>"MyHobbyStore Retail", "Seller_price_1"=>"??9.60"}, {"Seller_name_2"=>"Gaugemaster", "Seller_price_2"=>"??14.50"}]]}

MyHobbyStore Retail,Gaugemaster

5) **Tokenize column 'product_name' for recommendation system**
   - by using the functions below:

```
import nltk
from nltk.tokenize import sent_tokenize, word_tokenize
from nltk.corpus import stopwords
from string import punctuation
```

product_name

S-60 EF66 Electric Locomotive (Tomica PlaRail Model Train) (japan import)

product_name_tokens

S-60,EF66,Electric,Locomotive,Tomica,PlaRail,Model,Train,japan,import

Data Preprocessing

6) **Tokenize column "Product Information" for recommendation system**

7) **Sentence Tokenize column "customer reviews" for sentiment analysis**

Technical Details Item Weight640 g Product Dimensions29.6 x 20.8 x 1 cm Manufacturer recommended age:6 years and up Item model numberR8148 Main Language(s)English manual, English Number of Game Players1 Number of Puzzle Pieces1 Assembly RequiredNo Scale1:72 Engine Typeelectric Track Width/GaugeHO Batteries Required?No Batteries Included?No Material Type(s)Paper Material Care InstructionsNo Remote Control Included?No Radio Control Suitabilityindoor Colorwhite ???? Additional Information ASINB00HJ208KO Best Sellers Rank 52,854 in Toys & Games (See top 100) #69 in??Toys & Games > Model Trains & Railway Sets > Rail Vehicles > Trains Shipping Weight640 g Delivery Destinations:Visit the Delivery Destinations Help page to see where this item can be delivered. Date First Available24 Dec. 2013 ???? Feedback ??Would you like to update product info or give feedback on images?

['Technical',
'Details',
'Item',
'Weight640',
'Product',
'Dimensions29.6',
'cm',
'Manufacturer',
'recommended',
'age:6',
'years',
'Item',
'model',
'numberR8148',
'Main',
'Language',
'English',
'manual',
'English',
'Number',
'Game',
'Players1',
'Number',
'Puzzle',
'Pieces1',

great model at great price // 5.0 // 17 April 2014 // By

Rayzalaff

on 17 April 2014 // This model is something all BR modellers will want and comes at a very reasonable price for a DCC fitted loco. will definitely recommend it & the seller | Fantastic model of a locomotive I've followed from childhood // 4.0 // 21 Feb. 2015 // By

Chris

on 21 Feb. 2015 // Fantastic model of a locomotive I've followed from childhood. NYMR had a fabulous variant steaming and it's a joy to behold. | Five Stars // 5.0 // 5 May 2015 // By

PAUL STICKLAND

on 5 May 2015 // FANTASTIC DETAIL, HAVE NOT RUN IT YET. BUILD LAYOUT PLANS AT THE AT THIS TIME.

['great model at great price ',
" This model is something all BR modellers will want and comes at a very reasonable price for a DCC fitted loco. will definitely recommend it & the seller , Fantastic model of a locomotive I've followed from childhood ",
" Fantastic model of a locomotive I've followed from childhood. NYMR had a fabulous variant steaming and it's a joy to behold. , Five Stars ",
' FANTASTIC DETAIL, HAVE NOT RUN IT YET. BUILD LAYOUT PLANS AT THE AT THIS TIME.']

Sentiment Analysis

1) **Use**

```
from nltk.sentiment.vader import SentimentIntensityAnalyzer
```

To analyse the sentence tokens of each product

The threshold of the compound scores are set in line with the standard suggested by the documentation

It is also useful for researchers who would like to set standardized thresholds for classifying sentences as either positive, neutral, or negative. Typical threshold values (used in the literature cited on this page) are:

1. positive sentiment: `compound` score >= 0.05
2. neutral sentiment: (`compound` score > -0.05) and (`compound` score < 0.05)
3. negative sentiment: `compound` score <= -0.05

```
positive     8641
neutral      1057
negative      302
Name: summary_sentiment, dtype: int64
```

['great model at great price ',
 " This model is something all BR modellers will want and comes at a very reasonable price for a DCC fitted loco. will definitely recommend it & the seller , Fantastic model of a locomotive I've followed from childhood ",
 " Fantastic model of a locomotive I've followed from childhood. NYMR had a fabulous variant steaming and it's a joy to behold. , Five Stars ",
 ' FANTASTIC DETAIL, HAVE NOT RUN IT YET. BUILD LAYOUT PLANS AT THE AT THIS TIME.']

```
def sentiment_analyzer(review_col):
    """
    input the review series, output a new series named
    'summary_sentiment' in where 'positive / neutral / negative'
    """
    summary_sentiment = []
    sia = SentimentIntensityAnalyzer()
    row = 0
    for product in review_col:
        for sentence in product:
            sum_compound = 0
            ss = sia.polarity_scores(sentence)
            sum_compound += ss['compound']
        avg_compound = sum_compound/(len(review_col[row]))

        if avg_compound >= 0.05:
            summary_sentiment.append('positive')
        elif (avg_compound > -0.05) and (avg_compound < 0.05):
            summary_sentiment.append('neutral')
        else:
            summary_sentiment.append('negative')
        row += 1
    return summary_sentiment
```

03

Content-based Recommendation System

Feature Selection

```python
def combine_tokens(clean_comb_df):
    combined_tokens = []
    for i in range(clean_comb_df.shape[0]):
        empty = ''
        empty += clean_comb_df['product_name_tokens'][i]
        empty += clean_comb_df['amazon_category_and_sub_category'][i]
        empty += clean_comb_df['price_tag'][i]
        empty += clean_comb_df['seller_list_tokens'][i]
        empty += clean_comb_df['product_info_token'][i]

        combined_tokens.append(empty)
    clean_comb_df['combined_tokens'] = combined_tokens
    return clean_comb_df
```

Cosine Similarity

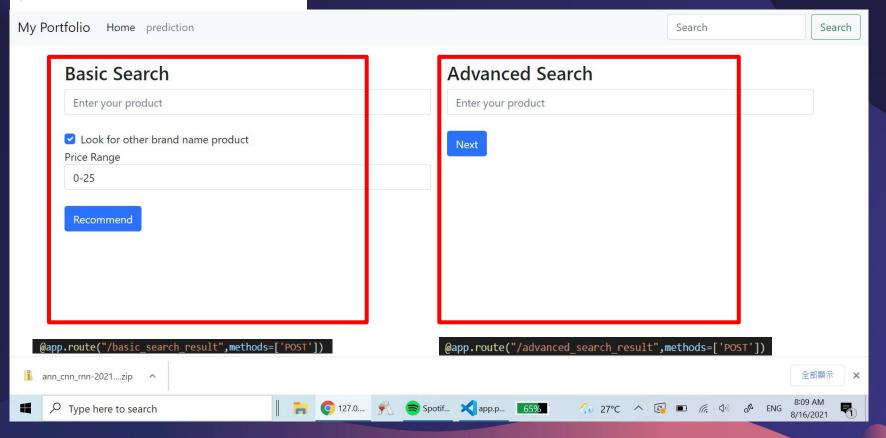-Using the attributes of the products to generate a cosine similarity matrix

```
cos = cosine_similarity(X)
print(cos)
```

```
[[1.         0.22562957 0.28425352 ... 0.03909948 0.05517062 0.10491307]
 [0.22562957 1.         0.31750324 ... 0.02674593 0.03316357 0.02895974]
 [0.28425352 0.31750324 1.         ... 0.03959597 0.04434981 0.05601553]
 ...
 [0.03909948 0.02674593 0.03959597 ... 1.         0.05822906 0.05527752]
 [0.05517062 0.03316357 0.04434981 ... 0.05822906 1.         0.0564913 ]
 [0.10491307 0.02895974 0.05601553 ... 0.05527752 0.0564913  1.        ]]
```

```
print(cos.shape)
```

```
(10000, 10000)
```

Flask

- Micro framework for Python

- Web Application

- Jinja Templating {{ }} / {% %}

prediction.html

Search   Search

Basic Search

Enter your product

☑ Look for other brand name product

Price Range

0-25

Recommend

Advanced Search

Enter your product

Next

```python
@app.route("/basic_search_result",methods=['POST'])
```

```python
@app.route("/advanced_search_result",methods=['POST'])
```

ann_cnn_rnn-2021....zip   全部顯示   ✕

Type here to search   127.0....   Spotif...   app.p...   65%   27°C   ENG   8:09 AM   8/16/2021

```
<option selected>Choose desired product to sell</option>
{% for selection in selections%}
<option >
{{selection}}
</option>
{% endfor %}
```

My Portfolio   Home   prediction

Keywords: tree

Filtered Product
Choose desired product to sell

Choose desired product to sell

20pcs Scale 1/200 Model Single Head Street Lights Lamppost

10pcs Scale 1/500 3V Single-Head Model Street Lights Lamppost for Scenary Lighting

Sesame Street - Snuffleupagus 17"

Shadowrun Street Samurai Dice Set by Q-workshop

Playmobil 5496 Advent Calendar Christmas Room with Tree

Playmobil 3943 - Christmas Tree Decorations

Snowman Forest Animals Decorating the Christmas Tree Advent Calendar 24 x20 cm

Takaratomy 819738 Official Pokemon X and Y MC-056 2" Treecko Action Figure

A Nightmare on Elm Street - Freddy Krueger Bobble Head

Paint By Numbers Kits Canvas Mounted on Wood Frame with Brushes and Paints by *Moon's Outlet* Reverside House + Girl + Hug + Plum Blossom + Love of Birds + Po

ann_c

Type here to search

27°C   ENG   8:20 AM
                  8/16/2021

My Portfolio   Home   prediction

Search   Search

Here is the top 10 recommended product for sale

Items	Manufacturer	Price Range	Sentiment	Link
20pcs Model Street Lights Lamppost w/ Single Head Scale 1/75	Generic	0-25	positive	Click me
20pcs Model Garden Light Double Heads Lamppost Scale 1:100	Generic	0-25	positive	Click me
10pcs Scale 1/500 3V Single-Head Model Street Lights Lamppost for Scenary Lighting	Generic	0-25	positive	Click me
20Pcs Single-Head Model Garden Lights Lamppost Lamp Layout Scale: 1/100--Black	Generic	0-25	positive	Click me
I/G (2) Oldtime Lanterns Working	Busch	50-100	positive	Click me
Topro LED Flashing Neon Sensory Maracas Light Up Maracas Toy	Topro	No Price	positive	Click me
Viessmann 64706 H0 6pc set Park lamps	Viessmann	No Price	positive	Click me
Viessmann 4721	Viessmann	No Price	positive	Click me
Train Tech SK2 Home Signal Kit with R/G LEDs	Train Tech	0-25	positive	Click me
Gaugemaster HBYS Hornby Type Colour Light Signal for 12v DC	Gaugemaster	0-25	positive	Click me

20Pcs Single-Head Model Garden Lights Lamppost Lamp Layout Scale: 1/100--Black

Challenges

05

Data Cleaning

-Remove unnecessary information from database is not easy

-Our database contains 10 thousand rows and 15 columns

Feature selection

Recommender system

-With capturing more of the finer details, the better our recommender system perform.

-e.g. product name, category, price, seller, product information

Challenges

Flask

Encounter problem with creating Brand/Category drop down list in website for user to choose

Encounter problem with embedding Wordcloud in website for each recommendation

Image

Provide support with image recognition function

Collaborative Recommendation

Upgrade our searching technique to provides more accurate result

Search Function

Provide support on adding more searching criteria, e.g. Brand

Data Base

Collect customers demographic information in accordance with the legal requirements

Q & A

THANKS!