Restaurant recommendation system with Python

April 4, 2021

To create the Restaurant recommendation system, I will create a content-based recommendation system where when I enter the name of a restaurant, the Restaurant recommendation system will look at reviews from other restaurants, and System will recommend us to the other restaurants with similar reviews and sort them from the top-rated.

```
[5]: import pandas as pd
     import numpy as np
     import seaborn as sns
     import matplotlib.pyplot as plt
     import re
     import warnings
     warnings.filterwarnings('always')
     warnings.filterwarnings('ignore')
     from sklearn.metrics.pairwise import linear_kernel
     from sklearn.metrics.pairwise import cosine_similarity
     from sklearn.feature_extraction.text import CountVectorizer
     from sklearn.feature_extraction.text import TfidfVectorizer
[6]: df=pd.read_csv('zomato.csv')
[7]: df.head()
[7]:
                                                       url \
     0 https://www.zomato.com/bangalore/jalsa-banasha...
     1 https://www.zomato.com/bangalore/spice-elephan...
     2 https://www.zomato.com/SanchurroBangalore?cont...
     3 https://www.zomato.com/bangalore/addhuri-udupi...
     4 https://www.zomato.com/bangalore/grand-village...
                                                                             name
     0 942, 21st Main Road, 2nd Stage, Banashankari, ...
                                                                          Jalsa
     1 2nd Floor, 80 Feet Road, Near Big Bazaar, 6th ...
                                                                 Spice Elephant
     2 1112, Next to KIMS Medical College, 17th Cross...
                                                                San Churro Cafe
     3 1st Floor, Annakuteera, 3rd Stage, Banashankar...
                                                         Addhuri Udupi Bhojana
     4 10, 3rd Floor, Lakshmi Associates, Gandhi Baza...
                                                                  Grand Village
       online_order book_table
                                 rate votes
                                                                          phone \
```

```
0
                Yes
                           Yes 4.1/5
                                          775
                                                 080 42297555\r\n+91 9743772233
                                4.1/5
                                          787
     1
                Yes
                            No
                                                                    080 41714161
     2
                Yes
                            No
                                3.8/5
                                          918
                                                                 +91 9663487993
     3
                                3.7/5
                 No
                            No
                                           88
                                                                  +91 9620009302
                 No
                                3.8/5
                                          166
                                               +91 8026612447\r\n+91 9901210005
                            No
            location
                                rest_type \
        Banashankari
                            Casual Dining
     1 Banashankari
                            Casual Dining
     2 Banashankari
                      Cafe, Casual Dining
     3 Banashankari
                              Quick Bites
     4 Basavanagudi
                            Casual Dining
                                                dish_liked \
     O Pasta, Lunch Buffet, Masala Papad, Paneer Laja...
     1 Momos, Lunch Buffet, Chocolate Nirvana, Thai G...
        Churros, Cannelloni, Minestrone Soup, Hot Choc...
     3
                                               Masala Dosa
     4
                                       Panipuri, Gol Gappe
                              cuisines approx_cost(for two people)
        North Indian, Mughlai, Chinese
                                                                 800
     0
     1
           Chinese, North Indian, Thai
                                                                 800
     2
                Cafe, Mexican, Italian
                                                                800
     3
            South Indian, North Indian
                                                                 300
              North Indian, Rajasthani
                                                                 600
                                              reviews_list menu_item
     O [('Rated 4.0', 'RATED\n A beautiful place to ...
                                                                 1 [('Rated 4.0', 'RATED\n Had been here for din...
                                                                 2 [('Rated 3.0', "RATED\n
                                 Ambience is not that ...
                                                                 3 [('Rated 4.0', "RATED\n
                                 Great food and proper...
                                                                 4 [('Rated 4.0', 'RATED\n
                                Very good restaurant ...
                                                                 listed_in(type) listed_in(city)
     0
                Buffet
                          Banashankari
                Buffet
     1
                          Banashankari
     2
                Buffet
                          Banashankari
     3
                Buffet
                          Banashankari
     4
                Buffet
                          Banashankari
[8]: df.shape
[8]: (51717, 17)
     df.columns
```

```
[9]: Index(['url', 'address', 'name', 'online_order', 'book_table', 'rate', 'votes',
             'phone', 'location', 'rest_type', 'dish_liked', 'cuisines',
             'approx_cost(for two people)', 'reviews_list', 'menu_item',
             'listed_in(type)', 'listed_in(city)'],
            dtype='object')
[10]: zomato=df.drop(['url','dish_liked','phone'],axis=1)
[11]: zomato.duplicated().sum()
[11]: 43
[12]: zomato.drop_duplicates(inplace=True)
[13]: zomato.duplicated().sum()
[13]: 0
[14]: zomato.isnull().sum()
[14]: address
                                         0
                                         0
      name
      online_order
                                         0
      book_table
                                         0
                                      7767
      rate
      votes
                                        21
      location
     rest_type
                                       227
      cuisines
                                        45
      approx_cost(for two people)
                                       345
      reviews_list
                                         0
     menu item
                                         0
      listed_in(type)
                                         0
      listed_in(city)
      dtype: int64
[15]: zomato.dropna(inplace=True)
[16]: zomato.isnull().sum()
[16]: address
                                      0
                                      0
      name
      online_order
                                      0
      book_table
                                      0
      rate
                                      0
                                      0
      votes
                                      0
      location
                                      0
      rest_type
```

```
cuisines
                                   0
     approx_cost(for two people)
                                   0
     reviews_list
                                   0
                                   0
     menu_item
     listed_in(type)
                                   0
     listed_in(city)
                                   0
     dtype: int64
[17]: zomato=zomato.rename(columns={'approx_cost(for two people)':
      [18]: zomato.columns
[18]: Index(['address', 'name', 'online_order', 'book_table', 'rate', 'votes',
            'location', 'rest_type', 'cuisines', 'cost', 'reviews_list',
            'menu_item', 'type', 'city'],
           dtype='object')
[19]: zomato['cost']=zomato['cost'].astype(str)
     zomato['cost'] = zomato['cost'].apply(lambda x:x.replace(',','.'))
     zomato['cost'] = zomato['cost'].astype(float)
     zomato['cost']
[19]: 0
              0.008
              800.0
     1
     2
              800.0
     3
              300.0
              600.0
     51709
              800.0
              0.008
     51711
     51712
                1.5
     51715
                2.5
     51716
                1.5
     Name: cost, Length: 43499, dtype: float64
[20]: zomato['rate']
[20]: 0
               4.1/5
               4.1/5
     1
     2
               3.8/5
     3
               3.7/5
               3.8/5
     51709
              3.7 /5
     51711
              2.5 /5
              3.6 /5
     51712
```

```
3.4 /5
      51716
      Name: rate, Length: 43499, dtype: object
[21]: #Removing '/5' from ratings
      zomato= zomato.loc[zomato.rate !='NEW']
      zomato=zomato.loc[zomato.rate !='-'].reset index(drop=True)
      remove_slash = lambda x:x.replace('/5','') if type(x)==np.str else x
      zomato.rate= zomato.rate.apply(remove slash).str.strip().astype('float')
[22]: #djusting the column names
      zomato.name=zomato.name.apply(lambda x:x.title())
      zomato.online_order.replace(('Yes','No'),('True','False'),inplace=True)
      zomato.book table.replace(('Yes','No'),('True','False'),inplace=True)
[23]: zomato.head(5)
[23]:
                                                   address
                                                                              name
      0 942, 21st Main Road, 2nd Stage, Banashankari, ...
                                                                           Jalsa
      1 2nd Floor, 80 Feet Road, Near Big Bazaar, 6th ...
                                                                  Spice Elephant
      2 1112, Next to KIMS Medical College, 17th Cross...
                                                                 San Churro Cafe
      3 1st Floor, Annakuteera, 3rd Stage, Banashankar... Addhuri Udupi Bhojana
      4 10, 3rd Floor, Lakshmi Associates, Gandhi Baza...
                                                                   Grand Village
        online order book table
                                                  location
                                                                       rest_type \
                                 rate
                                       votes
      0
                True
                           True
                                  4.1
                                         775 Banashankari
                                                                   Casual Dining
      1
                True
                          False
                                         787 Banashankari
                                                                   Casual Dining
                                  4.1
      2
                                         918 Banashankari Cafe, Casual Dining
                True
                          False
                                  3.8
      3
               False
                          False
                                              Banashankari
                                                                     Quick Bites
                                  3.7
                                          88
               False
      4
                          False
                                  3.8
                                         166 Basavanagudi
                                                                   Casual Dining
                               cuisines
                                          cost
         North Indian, Mughlai, Chinese
                                         800.0
      1
            Chinese, North Indian, Thai
                                         800.0
      2
                 Cafe, Mexican, Italian
                                         800.0
      3
             South Indian, North Indian
                                         300.0
      4
               North Indian, Rajasthani
                                         600.0
                                              reviews_list menu_item
                                                                         type \
      O [('Rated 4.0', 'RATED\n A beautiful place to ...
                                                                 [] Buffet
      1 [('Rated 4.0', 'RATED\n Had been here for din...
                                                                 [] Buffet
      2 [('Rated 3.0', "RATED\n Ambience is not that ...
                                                                 [] Buffet
      3 [('Rated 4.0', "RATED\n Great food and proper...
                                                                 [] Buffet
      4 [('Rated 4.0', 'RATED\n Very good restaurant ...
                                                                 [] Buffet
                 city
```

51715

4.3 / 5

Banashankari

```
2 Banashankari
     3 Banashankari
     4 Banashankari
[24]: #Computing Mean rating
     restaurants=list(zomato['name'].unique())
     zomato['Mean_rating']=0
     for i in range(len(restaurants)):
      \rightarrowmean()
[25]: from sklearn.preprocessing import MinMaxScaler
     scaler=MinMaxScaler(feature_range=(1,5))
     zomato[['Mean_rating']]=scaler.fit_transform(zomato[['Mean_rating']]).round(2)
[26]: #Lower casing
     zomato['reviews_list'] = zomato['reviews_list'].str.lower()
[27]: import string
     PUNCT_TO_REMOVE=string.punctuation
     def remove_punctuation(text):
         return text.translate(str.maketrans('','',PUNCT_TO_REMOVE))
     zomato['reviews_list']=zomato['reviews_list'].apply(lambda text:__
      →remove_punctuation(text))
[28]: import nltk
     nltk.download('stopwords')
     [nltk_data] Downloading package stopwords to
     [nltk_data]
                    C:\Users\admin\AppData\Roaming\nltk_data...
     [nltk_data]
                  Package stopwords is already up-to-date!
[28]: True
[29]: #Removal of stopwords
     from nltk.corpus import stopwords
     STOPWORDS= set(stopwords.words('english'))
     def remove_stopwords(text):
         return ''.join([word for word in str(text).split() if word not in_
      →STOPWORDS])
     zomato['reviews_list'] = zomato['reviews_list'].apply(lambda text:
      →remove_stopwords(text))
```

1 Banashankari

```
[30]: #Removal of Urls
      def remove_urls(text):
          url_pattern=re.compile(r'https?://\S+|www\.\S+')
          return url_pattern.sub(r'',text)
      zomato['reviews_list'] = zomato['reviews_list'].apply(lambda text:
       →remove urls(text))
[31]: zomato[['reviews list', 'cuisines']].sample(10)
[31]:
                                                   reviews_list \
      40107
             rated40ratednwentsaturdaylunchorderedvegchicke...
      15040
             rated50ratednonebestindoripohasbangalorendefin...
      6098
             rated50ratednfoodtasteyummyqualityquantitywort...
      26265 rated40ratednplacenarrowstreetindiranagarfishc...
      8872
             rated50ratednutterlydeliciousbengalifoodordere...
      39280
             rated40ratedngoodfoodbudgetorderedhomedelivery...
      792
             rated20ratedn15tbhgobichillmushroomchilliuptom...
      32110
      30530 rated50ratedn5starsnamequalityquantityicecream...
      9547
             rated40ratednhealthyfreshprovidegoodoptionsple...
                                           cuisines
      40107
                      Biryani, North Indian, Kebab
      15040
                            Street Food, Fast Food
      6098
                      North Indian, Bengali, Oriya
      26265
             Oriya, North Indian, Bengali, Chinese
      8872
                             Bengali, North Indian
      39280
                                           Lebanese
      792
                             Chinese, North Indian
      32110
                                       South Indian
      30530
                    Desserts, Fast Food, Beverages
      9547
                                Mexican, Fast Food
[32]: ##Restaurant Names
      restaurant_names = list(zomato['name'].unique())
      def get_top_words(column,top_nu_of_words,nu_of_word):
          vec=CountVectorizer(ngram_range=nu_of_word,stop_words='english')
          bag_of_words=vec.fit_transform(column)
          sum words=bag of words.sum(axis=0)
          words_freq=[(word,sum_words[0,idx]) for word,idx in vec.vocabulary_.items()]
          words_freq=sorted(words_freq,key=lambda x:x[1],reverse=True)
          return words_freq[:top_nu_of_words]
[33]: | zomato=zomato.drop(['address', 'rest_type', 'type', 'menu_item', 'votes'], axis=1)
```

```
[34]: df_percent= zomato.sample(frac=0.5)
[35]: df_percent.set_index('name',inplace=True)
    df_percent
[35]:
                                              online_order book_table \
    name
    55 Wall Street
                                                    False
                                                              True
    Chili'S American Grill & Bar
                                                     True
                                                              True
    The Terminus - The Oterra
                                                    False
                                                              True
    Shanthi Sagar
                                                     True
                                                             False
    Hotel Salala
                                                             False
                                                     True
    Hotel Devanand Kerala Restaurant
                                                             False
                                                     True
    True
                                                   False
    Foxtrot - House Of Subculture
                                                    False
                                                              True
    False
                                                   False
    I-Bar - The Park Bangalore
                                                    False
                                                             False
                                               rate \
    name
    55 Wall Street
                                                3.9
    Chili'S American Grill & Bar
                                                4.6
    The Terminus - The Oterra
                                                4.0
    Shanthi Sagar
                                                4.0
    Hotel Salala
                                                3.6
    Hotel Devanand Kerala Restaurant
                                                3.4
    4.1
    Foxtrot - House Of Subculture
                                                4.4
    3.2
    I-Bar - The Park Bangalore
                                                3.8
                                                          location \
    name
                                                          JP Nagar
    55 Wall Street
    Chili'S American Grill & Bar
                                                         Bellandur
    The Terminus - The Oterra
                                                    Electronic City
    Shanthi Sagar
                                                            Ulsoor
    Hotel Salala
                                                        Brookefield
    Hotel Devanand Kerala Restaurant
                                                     CV Raman Nagar
    Richmond Road
    Foxtrot - House Of Subculture
                                               Koramangala 3rd Block
    MG Road
    I-Bar - The Park Bangalore
```

cuisines \

name 55 Wall Street Finger Food, North Indian, Chinese Chili'S American Grill & Bar American, Tex-Mex, Burger, BBQ The Terminus - The Oterra Finger Food, Continental Shanthi Sagar South Indian, North Indian, Chinese, Juices Hotel Salala North Indian, South Indian, Chinese, Biryani Hotel Devanand Kerala Restaurant Kerala Wraps, Pizza, Cafe Foxtrot - House Of Subculture Cafe, American, Asian, North Indian Cafe, American, European I-Bar - The Park Bangalore North Indian, Chinese, Mediterranean cost \ name 55 Wall Street 1.3 Chili'S American Grill & Bar 1.8 The Terminus - The Oterra 2.5 Shanthi Sagar 250.0 Hotel Salala 400.0 Hotel Devanand Kerala Restaurant 300.0 Foxtrot - House Of Subculture 1.0 I-Bar - The Park Bangalore 3.0 reviews_list \ name

55 Wall Street rated50ratedntwowaysenjoyonefloordiscenjoylive...

Chili'S American Grill & Bar

rated40ratedngoodtimesspandherenwecomefriendsf...

The Terminus - The Oterra

rated50ratednexcellentplacebitfarmaincityuvisi...

	city `
name	
55 Wall Street	BTM
Chili'S American Grill & Bar	Sarjapur Road
The Terminus - The Oterra	Electronic City
Shanthi Sagar	Residency Road
Hotel Salala	Marathahalli
Hotel Devanand Kerala Restaurant	Old Airport Road
Cafã Â Ã Â Ã Â Â Â Â Â Â Â Â Â Â Â Â Â Â Â	Lavelle Road
Foxtrot - House Of Subculture	Koramangala 5th Block
Cafã Â Ã Â Â Â Â Â Â Â Â Â Â Â Â Â Â Â Â Â	Koramangala 7th Block
I-Bar - The Park Bangalore	MG Road

	Mean_rati	ng
name		
55 Wall Street	3.	71
Chili'S American Grill & Bar	4.	70
The Terminus - The Oterra	3.	84
Shanthi Sagar	3.	44
Hotel Salala	3.	39
	•••	
Hotel Devanand Kerala Restaurant	3.	06
Cafã Â Ã Â Ã Â Ã Â Ã Â Ã Â Ã Â Ã Â Â Â Â Â	3.97	
Foxtrot - House Of Subculture	4.	35
Cafã Â Ã Â Â Â Â Â Â Â Â Â Â Â Â Â Â Â Â Â	2.81	
I-Bar - The Park Bangalore	3.	58

[20618 rows x 9 columns]

```
[36]: indices=pd.Series(df_percent.index)
     indices
[36]: 0
                                              55 Wall Street
     1
                                 Chili'S American Grill & Bar
     2
                                    The Terminus - The Oterra
     3
                                              Shanthi Sagar
     4
                                               Hotel Salala
     20613
                             Hotel Devanand Kerala Restaurant
     20614
             Foxtrot - House Of Subculture
     20615
     20616
             20617
                                   I-Bar - The Park Bangalore
     Name: name, Length: 20618, dtype: object
[]: from sklearn.metrics.pairwise import linear_kernel
     from sklearn.feature_extraction.text import CountVectorizer
     from sklearn.feature_extraction.text import TfidfVectorizer
     #Creating Tf-IDF matrix
     tfidf = TfidfVectorizer(analyzer='word', ngram_range=(1, 2), min_df=0,__
      ⇔stop_words='english')
     tfidf_matrix=tfidf.fit_transform(df_percent['reviews_list'])
     cosine_similarities=linear_kernel(tfidf_matrix,tfidf_matrix)
```

#Now the last step for creating a Restaurant Recommendation System is to write a function that will recommend restaurants:

```
for each in top30_indexes:
       recommend_restaurant.append(list(df_percent.index)[each])
   # Creating the new data set to show similar restaurants
   df_new = pd.DataFrame(columns=['cuisines', 'Mean Rating', 'cost'])
   # Create the top 30 similar restaurants with some of their columns
   for each in recommend_restaurant:
       df_new = df_new.append(pd.DataFrame(df_percent[['cuisines', 'Mean_
→Rating', 'cost']][df_percent.index == each].sample()))
   # Drop the same named restaurants and sort only the top 10 by the highest \Box
\hookrightarrow rating
   df_new = df_new.drop_duplicates(subset=['cuisines', 'Mean Rating', 'cost'], u
→keep=False)
   df_new = df_new.sort_values(by='Mean Rating', ascending=False).head(10)
   print('TOP %s RESTAURANTS LIKE %s WITH SIMILAR REVIEWS: ' %_
return df_new
#recommend('Pai Vihar')
```

[]: