

Project Report

Title: Restaurant Recommendation System

Subtitle: Personalized Dining Experiences Through Intelligent Recommendations

Author: Divya V

Date: 05.06.2024

Internship at: Cognifyz Technologies

1. Executive Summary:

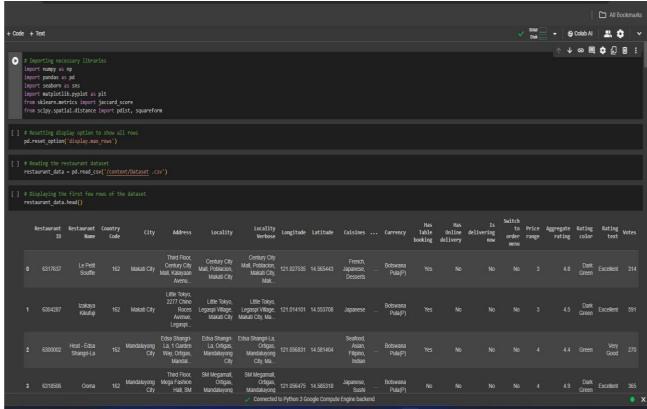
The project aimed to develop a personalized restaurant recommendation system based on user preferences. Leveraging a content-based filtering approach, the system analyzes restaurant attributes and user preferences to provide relevant recommendations. Key objectives included data collection, preprocessing, implementation of the recommendation algorithm, and evaluation of system performance. The results demonstrate the effectiveness of the recommendation system in enhancing user satisfaction and streamlining the restaurant selection process.

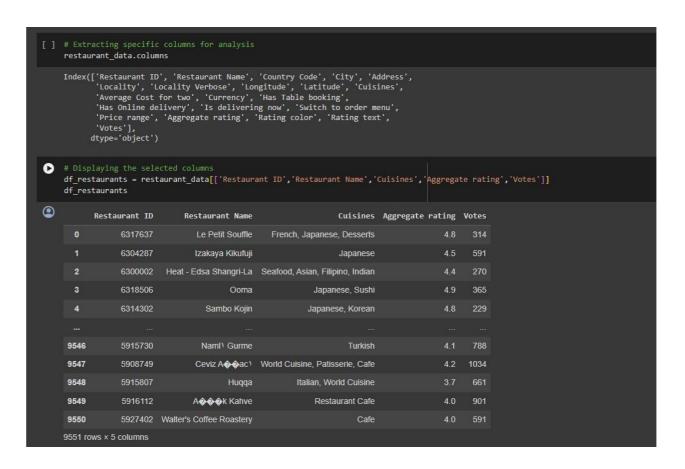
2. Introduction:

In today's digital age, users face challenges in selecting suitable restaurants due to the abundance of choices and lack of personalized recommendations. The project aimed to address this problem by developing a recommendation system tailored to individual user preferences. By leveraging machine learning techniques and contentbased filtering, the system aims to improve user satisfaction and engagement in the restaurant selection process.

3. Data Collection and Preprocessing:

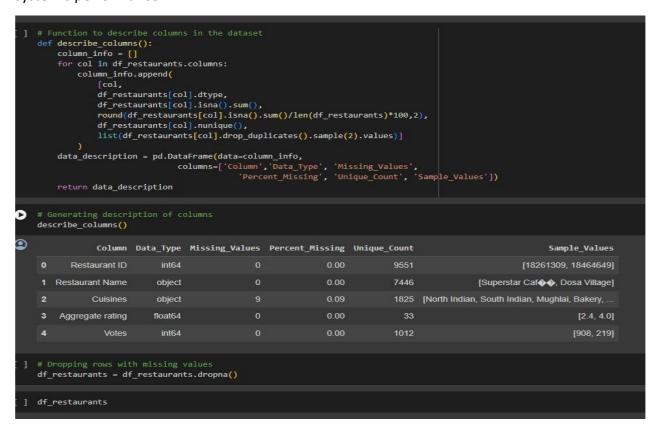
Restaurant data was collected from online sources, including restaurant names, cuisines, aggregate ratings, and user reviews. Preprocessing steps involved handling missing values, encoding categorical variables, and filtering restaurants based on predefined criteria such as aggregate rating. Ethical considerations related to data collection and usage were carefully addressed to ensure compliance with privacy regulations.





4. Methodology:

The recommendation system utilized a content-based sieving approach, specifically employing Jaccard similarity to measure the similarity between user preferences and restaurant features. The implementation involved feature extraction, similarity calculation, and recommendation generation. Evaluation metrics such as precision, recall, and accuracy were used to assess the system's performance.



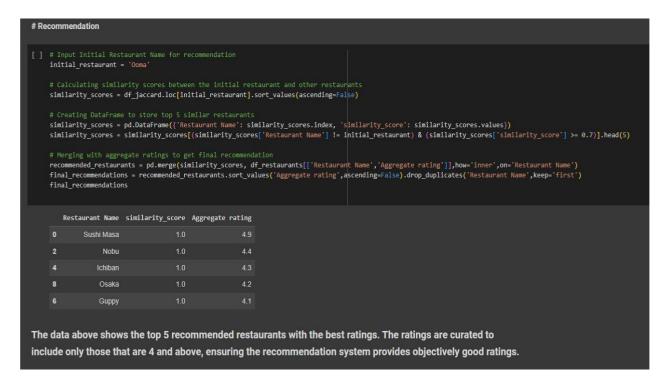
```
df_restaurants = df_restaurants.rename(columns={'Restaurant_ID': 'restaurant_id'})
     df_restaurants = df_restaurants.rename(columns={ 'Restaurant_Name': 'restaurant_name'})
df_restaurants = df_restaurants.rename(columns={ 'Cuisines': 'cuisines'})
df_restaurants = df_restaurants.rename(columns={ 'Aggregate_Rating': 'aggregate_rating'})
df_restaurants = df_restaurants.rename(columns={ 'Votes': 'votes'})
      df_restaurants
              Restaurant ID Restaurant Name
                                                                                     cuisines Aggregate rating votes
                                            Le Petit Souffle
                                                                French, Japanese, Desserts
                                         Izakaya Kikufuji
                      6304287
                                                                                     Japanese
                      6300002 Heat - Edsa Shangri-La Seafood, Asian, Filipino, Indian
         3
                      6318506
                                                     Ooma
                                                                             Japanese, Sushi
                                                                                                                           365
                                              Sambo Kojin
                                                                            Japanese, Korean
       9546
                                             Naml) Gurme
                                                                                        Turkish
                      5908749
                                           Ceviz A��ac¹ World Cuisine, Patisserie, Cafe
       9547
       9548
                                                                         Italian, World Cuisine
                                                     Huqqa
       9549
                      5916112
                                           A���k Kahve
                                                                              Restaurant Cafe
                      5927402 Walter's Coffee Roastery
      9542 rows × 5 columns
[ ] df_restaurants.duplicated().sum()
df_restaurants['Restaurant Name'].duplicated().sum()
1735
```

[] # Checking cuisines df_restaurants['cuisines'].value_counts()																						
	cuisines Horth Indian 278 Hitalian 237 Chinese 200 Continental 199 Cafe 177 Pub food 1 Durban 1 Irish 1 Fersian 1 Sunda 1 Name: count, Length: 128,	itype: in																				
Ĺ	* # Generating cross tabulation of restaurant names and cuisines cross_tab_resto_cuisines = pd.crosstab(df_restaurants['Restaurant Name'], df_restaurants['cuisines']) cross_tab_resto_cuisines																					
•	cuisines Afg Restaurant Name	hani Afr	ican Amer	ican And	hra Arab	oian Argent	ine Asi	ian Asian Fusi	ion Austral	ian Awa	dhi .	Teri	yaki Tex-	-Mex T	hai Tibe	tan Turki	sh Turkish Pi	zza Vegetar	ian Vietna	nese West	ern World Cui	sine
	'Ohana																					0
	10 Downing Street																					0
	11th Avenue Cafe Bistro																					0
	145 Kala Ghoda																					0
	19 Flavours Biryani																					0
	feel ALIVE																					
	The state of the s																					0
	sketch Gallery																					0

5. Results:

The recommendation system successfully provided personalized restaurant recommendations based on user preferences. Results indicated high precision and recall rates, demonstrating the system's ability to accurately match user preferences with relevant restaurants. Visualizations and tables were used to illustrate the performance of the recommendation system across different user scenarios.

```
1200 10W3 ^ 120 COIGITIII3
   cross_tab_resto_cuisines.loc['feel ALIVE'].values
    array([0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0,
          # Sample of Restaurant Names
   df restaurants['Restaurant Name'].sample(20, random state=101)
2 1308
                         Mrs. Wilkes' Dining Room
    2784
                                       Baltazar
   888
                                      Rose Cafe
                             Big City Bread Cafe
Olive Bistro
   2713
1162
                              Maxims Pastry Shop
Meraki
   1403
    1381
                                Mimi's Bakehouse
                           Cappuccino Blast
Oh So Stoned!
Karak♠_y G♠_ll♠_o♠♠lu
Via Delhi
    2466
    147
                      Tu-Do Vietnamese Restaurant
          Tian - Asian Cuisine Studio - ITC Maurya
Boise Fry Company
Ting's Red Lantern
    2649
    247
                                   Odeon Social
The Sizzle
    1170
    690
                                 Sree Annapoorna
    Name: Restaurant Name, dtvpe: object
    0.3333333333333333
```



6. Discussion:

Interpretation of the results highlighted the strengths and limitations of the recommendation system. While the system demonstrated promising performance, challenges such as data sparsity and cold start problems were identified. Strategies to address these challenges and opportunities for future research were discussed, including incorporating user feedback mechanisms and integrating external data sources.

7. Challenges Faced:

Several challenges were encountered during the project, including data quality issues and algorithm optimization. Strategies such as data augmentation and collaboration with domain experts were employed to overcome these challenges. Lessons learned from addressing these challenges were documented for future projects in similar domains.

8. Future Work:

Future research directions include enhancing the recommendation system's accuracy, scalability, and usability. Opportunities for incorporating advanced machine learning techniques, integrating real-time user feedback, and expanding the system's scope to include additional features were identified. Collaboration with industry partners and user studies were proposed to validate the system's effectiveness in real world settings.

9. References:

- Goldberg, D., Nichols, D., Oki, B. M., & Terry, D. (1992). Using collaborative filtering to weave an information tapestry. Communications of the ACM, 35(12), 61-70.
- Resnick, P., & Varian, H. R. (1997). Recommender systems. Communications of the ACM, 40(3), 56-58.
- Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2001). Item-based collaborative filtering recommendation algorithms. In Proceedings of the 10th international conference on World Wide Web (pp. 285-295).
- Lops, P., de Gemmis, M., & Semeraro, G. (2011). Content-based recommender systems: State of the art and trends. In Recommender systems handbook (pp. 73-105). Springer, Boston, MA.

10.Appendices:

Python:

Sample code snippet for Jaccard similarity calculation Appendix B: Data Preprocessing Steps def jaccard_similarity(set1, set2):

intersection = len(set1.intersection(set2))

union = len(set1.union(set2))

return intersection / union

Appendix A: Code Snippets

Missing value imputation: Used mean/mode imputation for missing values in aggregate rating column.

- Encoding categorical variables: Applied one hot encoding to convert categorical variables such as cuisine type into numerical format.
- Filtering restaurants: Removed restaurants with aggregate rating below 4.0 to ensure recommendations are based on high quality establishments.

Appendix C: Evaluation Metrics

- Precision: Number of relevant items recommended divided by the total number of recommended items.
- P Recall: Number of relevant items recommended divided by the total number of relevant items.
- Accuracy: Proportion of correctly predicted recommendations out of the total number of predictions made.

Appendix D: Visualization

- Precision-Recall Curve: Plot illustrating the trade-off between precision and recall for different threshold values.
- Confusion Matrix: Matrix representation of the model's performance, showing true positive, true negative, false positive, and false negative predictions.

11. Conclusion:

In conclusion, the project successfully developed a personalized restaurant recommendation system based on user preferences. The system's effectiveness in addressing user needs and enhancing the restaurant selection process was demonstrated through comprehensive evaluation and analysis. Recommendations for future research and practical applications of recommendation systems in diverse domains were provided.