Profiting Restaurants via Personalized Recommendation

Mehak Piplani, Shivam, Shubham Banka

1 Description and Innovation

The outbreak of COVID-19 has disrupted the food industry. There has been a decline in the number of people visiting restaurants. According to the study [1], the restaurant owners are facing issues in procuring stock for the dishes on the menu. The study also highlights the financial difficulty faced by the few operational restaurants. Our solution is to model a recommendation system that takes into account the cuisine and the locality for recommending new dishes for the menu. The most innovative aspect of our solution is the optimization of the ingredients supply chain by locating suppliers in COVID safe locations hence providing the restaurant owners a safe way to run their businesses cost-effectively.

2 Hypothesis

Our work focuses on helping restaurant owners by answering the following research questions:

- Is there a way to revamp the menu to avoid spending on low selling dishes? Is there a connection between profits and reduced menu size?
- Can we find a way to optimize the supply chain for the procurement of ingredients? Is it possible to include COVID safe locations?

We believe that profits are correlated to the number of items on the menu and the procurement method of ingredients required for those dishes. Hence, we focus on these two factors to evaluate if they can improve a restaurant owner's financial situation.

3 Literature Survey

The recipe recommendation system involves studying ingredient combination in recipes and food supply chains. The studies by authors [2, 3, 4] explain the rationale behind ingredient combination for different recipes. The studies [5, 6, 7] explain how recipes evolved and the studies [8, 9, 10] explain the effect of location nearness on recipe evolution. The studies [11, 12, 13] helped in understanding the food supply chain process. There is some work done on flavor combinations for various recipes using various computational tools [15, 16].

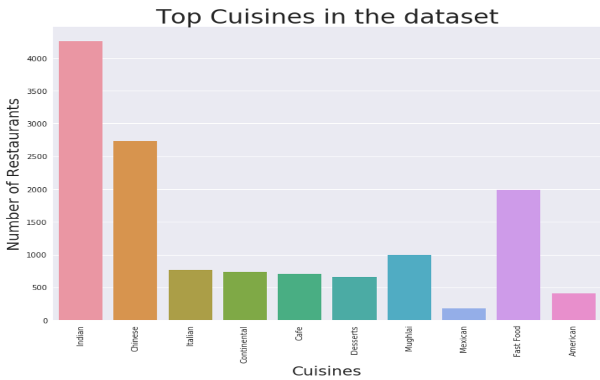


Figure 1: Frequency of Restaurants based on Cuisine

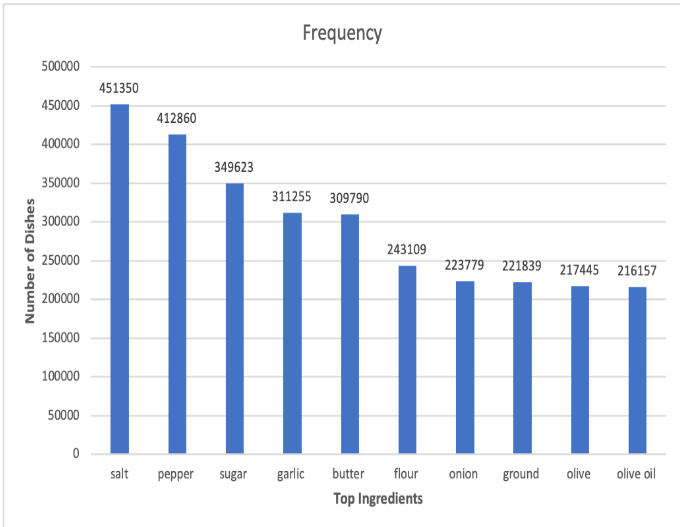


Figure 2: Top Ingredients

4 Dataset and Statistics

- **Frequently ordered dishes in the Neighborhood**
  To accomplish this task, data was scraped from Postmates's website. The dataset consists of the following parameters: **Restaurant Name, Rating, Frequently-ordered Items**. It was observed that Sushi items are most commonly ordered in the neighborhood of Los Angeles.

- **Dishes belonging to the same Cuisine**
  To make possible and validate the aforementioned task, data was collected through the Zomato API (world wide data) which provides details about the dishes from the top-rated restaurants serving the same cuisine.
  The dataset consists of 87k samples for the following parameters: **Restaurant Name, Location, Rating, Cuisine, Top-rated Items**. Fig 1 represents the distribution of the cuisines in the dataset via a bar graph

showcasing that the most frequent cuisine in the dataset is North Indian Cuisine and the least frequent is Mexican.

- **Dishes sharing the same ingredients**
  For this task, the dataset was collected from the Kaggle website.
  The dataset consists of 230186 unique samples for the following parameters: **Recipe Name, Ingredients, Minutes to prepare, Food.com tags for recipe**. Fig 2 represents the most frequent ingredients used in the dataset.

- **Information about suppliers**
  To accomplish this task, the United States Department of Agriculture (USDA) Organic Integrity Database which contains information for every organic product recognized by the USDA will be used.
  The dataset consists of 130373 samples for the following parameters: **Supplier Name, Address, Latitude, Longitude, ItemList**. Fig 3 shows the distribution of the different suppliers across the country.

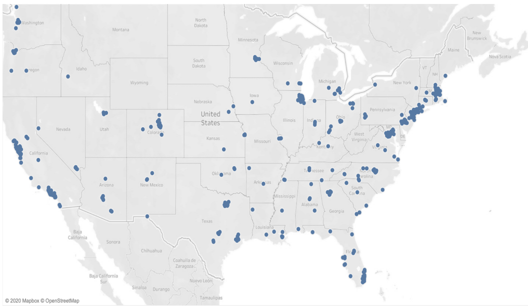


Figure 3: Supplier Locations as per USDA Organic Integrity Database

5 Method

5.1 Recipe Recommendation System

Our solution consists of building a Meta-level hybrid recommendation system. The input to the recommendation system includes the cuisine and location of the restaurant. The first level of the hybrid recommendation system takes the input and outputs the top five dishes for a restaurant. The first four dishes are extracted by applying the technique of content-based recommendation system on the zomato dataset profiling the restaurants based on the cuisines they serve. The fifth dish is extracted via an item-based recommendation technique applied to Postmates's dataset. For this, we extract the frequently ordered top-rated dishes in the neighborhood of the restaurant i.e. within five miles radius of the restaurant.
The second level of the hybrid recommendation model utilizes the dishes recommended by the first level. Ingredients are extracted for these dishes using the dataset. Since the names of the recommended dishes from the first level and the dataset do not match exactly, Levenshtein distance (Levenshtein distance between two words is the minimum number of single-character edits (insertions,

deletions or substitutions) required to change one word into the other) is used to calculate the similarity between the recommended dish name and the dish name from the dataset. These extracted ingredients are further used to find a set of five dishes that can be made from them when we have complete overlap with extracted ingredients.

5.2 Ingredient Supply Chain Optimization

The objective of this module is to minimize the operational cost of the procurement of ingredients. We use the list of ingredients required for cooking the recommended dishes and generate a list of suppliers selling these ingredients. For each ingredient, we find the nearest supplier by evaluating the euclidean distance between the location of the restaurant and the supplier. In case there are multiple ingredients available at any supplier, the distance of the restaurant from that supplier is added just once in the food miles. The module outputs a list of suppliers that have the minimum value of the Food Mile metric (the distance food travels from the supplier location to restaurant location).

6 Evaluation Metric

6.1 Recipe Recommendation System

We plan to evaluate the performance of our Meta-level hybrid recommendation model through an online survey since we do not have any ground truth data available. We conducted two surveys:

- The first survey asks the users to choose between a set of random five dishes v/s the set of five dishes recommended by our model for four sample restaurants.
- The second survey asks the users to rate the recommended dishes(top three) individually for four different restaurants on a scale of 0-5.

The rationale behind the survey is that the dishes which are rated highly have more chances of being ordered by people which in turn would increase the restaurant owner's profit.

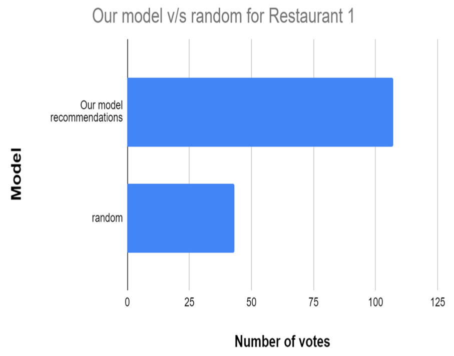
6.2 Ingredient Supply Chain Optimization

To evaluate the efficiency of the recommended supply chain, we compare our results to the worst-case scenario which procures ingredients from the locations farthest to the restaurant.

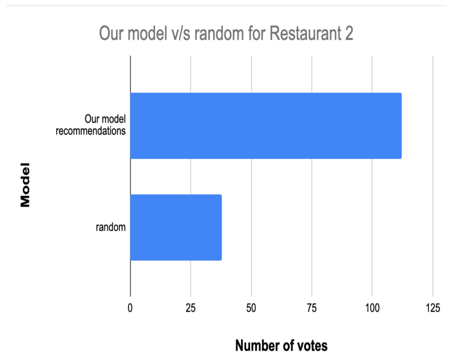
7 Results and Analysis

7.1 Recipe Recommendation System

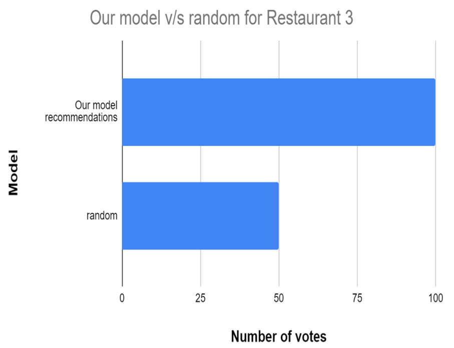
The results of the first survey conducted to evaluate the performance of the recommendation system are visualized in Fig 4. The results show that the set of 5 dishes recommended by our model has been chosen twice as much as the



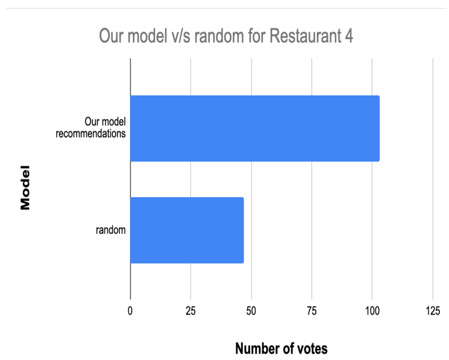
(a) Results for Restaurant 1



(b) Results for Restaurant 2



(c) Results for Restaurant 3



(d) Results for Restaurant 4

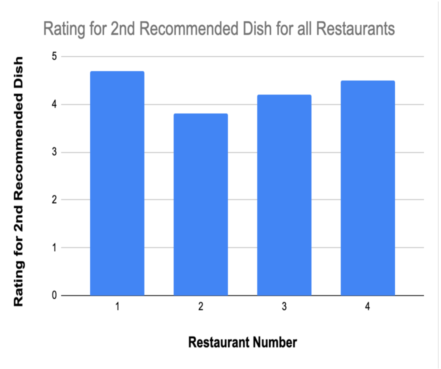
Figure 4: Survey result analysis

random set.
Fig 5 highlights the results of the second survey.
Fig 5a shows the average rating of the first recommended dish for each restaurant by 150 users signifying that the model recommends the first dish with 92% accuracy. The Fig 5b highlights the average rating of the second recommended dish for each restaurant by 150 users signifying that the model recommends the second dish with 86% accuracy. The Fig 5c highlights the average rating of the third recommended dish for each restaurant by 150 users signifying that the model recommends the third dish with 75% accuracy.
Overall, when combining the set of top five dishes recommendations, the model achieves an accuracy of 84.33%.

7.2 Ingredient Supply Chain Optimization

We observed an improvement of 46.5% on average in the food miles required for the procurement of ingredients. This will benefit the restaurant owners to get the ingredients faster and save a lot on the overall cost.

(a) Results for 1st recommended dish



(b) Results for 2nd recommended dish

(c) Results for 3rd recommended dish

Figure 5: Survey result analysis

8 Conclusion

The results in the above section signify that the model recommends dishes with approx 84.33% accuracy which are highly rated by the users. Since the users are liking our recommendations, it is highly likely they would order them and this would lead to an increase in profits. The supply chain of procurement of ingredients has been optimized and would lead to less spending. Hence, our belief about reduced menu size and optimized supply chain can help restaurant owners make more profit.

9 Future Work

Certain aspects of the project can be improved where we have faced difficulties. The ingredients data can be augmented with recipes consisting of all cuisines. We faced challenges with the inconsistencies in the name of the products at different locations. To tackle this, the product description can be used to identify similar products. The supply chain optimization can be extended to include the cost of the ingredient making it more robust and appropriate. Once completed, the project can be deployed as a web-app to open-source our findings.

References

[1] Song, Hyoung Yeon, Jihwan Lee, Seoki. (2021). "Impact of the COVID-19 Pandemic: Evidence from the U.S. Restaurant Industry" International Journal of Hospitality Management.

[2] Y.-Y. Ahn and S. Ahnert, "The Flavor Network" Leonard, vol. 46, no. 3, pp. 272–273, Mar. 2013

[3] Y.-Y. Ahn, S. E. Ahnert, J. P. Bagrow, and A.-L. Barabasi, "Flavor network and the principles of food pairing," Sci. Rep., vol. 1, Dec. 2011

[4] A. Bogojeska, S. Kalajdziski, and L. Kocarev, "Processing and Analysis of Macedonian Cuisine and its Flavours by Using Online Recipes," in ICT Innovations 2015, Springer, 2016, pp. 143–152.

[5] A. Jain and G. Bagler, "Culinary evolution models for Indian cuisines," Physica A: Statistical Mechanics and its Applications, vol. 503, pp. 170–176, 2018.

[6] O. Kinouchi, R. W. Diez-Garcia, A. J. Holanda, P. Zambianchi, and A. C. Roque, "The non-equilibrium nature of culinary evolution," New Journal of Physics, vol. 10, no. 7, p. 073020, 2008.

[7] Y. Kikuchi, M. Kumano, and M. Kimura, "Analyzing Dynamical Activities of Co-occurrence Patterns for Cooking Ingredients," in 2017 IEEE International Conference on Data Mining Workshops (ICDMW), 2017, pp. 17–24

[8] Y.-X. Zhu, J. Huang, Z.-K. Zhang, Q.-M. Zhang, T. Zhou, and Y.-Y. Ahn, "Geography and Similarity of Regional Cuisines in China," PLoS ONE, vol. 8, no. 11, p. e79161, Nov. 2013

[9] W. Min, B. K. Bao, S. Mei, Y. Zhu, Y. Rui, and S. Jiang, "You Are What You Eat: Exploring Rich Recipe Information for Cross-Region Food Analysis," IEEE Transactions on Multimedia, vol. PP, no. 99, pp. 1–1, 2017

[10] S. Sajadmanesh et al., "Kissing Cuisines: Exploring Worldwide Culinary Habits on the Web," in Proceedings of the 26th International Conference on World Wide Web Companion, Republic and Canton of Geneva, Switzerland, 2017, pp. 1013–1021

[11] L. Wang, Q. Li, N. Li, G. Dong, and Y. Yang, "Substructure Similarity Measurement in Chinese Recipes," in Proceedings of the 17th International Conference on World Wide Web, New York, NY, USA, 2008, pp. 979–988

[12] B. A. Wood et al., "Agricultural Science in the Wild: A Social Network Analysis of Farmer Knowledge Exchange," PLOS ONE, vol. 9, no. 8, p. e105203, Aug. 2014, doi: 10.1371/journal.pone.0105203. [26]A. Maertens and C. B. Barrett, "Measuring Social Networks' Effects on Agricultural Technology Adoption," Am J Agric Econ, vol. 95, no. 2, pp. 353–359, Jan. 2013

[13] IBM and Institute of Culinary Education, Cognitive Cooking with Chef Watson: Recipes for Innovation from IBM the Institute of Culinary Education. Naperville, Illinois: Sourcebooks, 2015.

[14] L. R. Varshney, F. Pinel, K. Varshney, D. Bhattacharjya, A. Schorgendorfer, and Y.-M. Chee, "A big data approach to computational creativity: The curious case of Chef Watson," IBM Journal of Research and Development, vol. 63, no. 1, pp. 7–1, 2019.

[15] N. Shao, P. Murali, and A. Sheopuri, "New Developments in Culinary Computational Creativity," in ICCC, 2014.

[16] F. Pinel, L. R. Varshney, and D. Bhattacharjya, "A Culinary Computational Creativity System," in Computational Creativity Research: Towards Creative Machines, T. R. Besold, M. Schorlemmer, and A. Smaill, Eds. Atlantis Press, 2015, pp. 327–346.