

# Yelp Rating Interpretation with Text-based and Graph-based Features

Zhuoran Liu, Mingye Chen

Dept. of Computer Science

Columbia University

zl2621@columbia.edu, mc4414@columbia.edu

**Abstract**—Interpreting and understanding how much impact different factors can have on users’ final ratings for restaurants is of key interest to restaurant owners who may later improve their businesses accordingly. Previously, researches on Yelp rating prediction focused heavily on pushing forward task-related metrics such as accuracy and RMSE. In this work, we propose to interpret the prediction process *per se* by building a random forest classifier for rating polarity prediction which rely only on transparent features from review texts and relation graphs of users and restaurants. Our classifier achieved reasonable accuracy, and provided meaningful insight into how users are evaluating restaurants.

**Keywords**—Interpretation; Yelp; graph; text; random forest

## I. INTRODUCTION

Rating prediction is one of the most important research topics in the field of recommendation systems. A successful rating prediction model should not only enable merchants to target a specific group of users, but also allow them to identify key factors that guarantees their success on specific products and targeted users. To this end, researchers incorporated as much contextual information into their prediction models in an attempt to capture potential factors that might play a key role. Reviews are one of the most commonly seen context being utilized, and a lot of sophisticated models have been developed to push the state-of-the-art performance to a satisfactory territory in terms of commercial use.

For these near-perfect models, however, there is one inherent problem that constantly haunts most of them - the lack of transparency. As a result, a merchant only knows that certain group of users are most attracted by certain products, but is agnostic of the reason behind this favoritism. This prevents merchants from sophisticating their strategies that caters to different user groups under various scenarios.

The lack of transparency and interpretability is mainly attributed to two reasons: First, the model itself is a black-box or too complicated to infer any intuitive information from; Second, the input into the model are features in some uninterpretable spaces that are projected by certain irreversible methods, e.g. matrix factorization, embedding methods.

To tackle this black-box issue, we propose to leverage only the most transparent algorithms and features. The algorithm we used in this work is random forest classifier,

which comes with a natural explanation for classification decision with regard to each feature. We selected the most naive yet clearly perceivable features based on n-grams from review texts, as well as graph features based on relations between users and business.

One potential challenge for building an interpretable rating prediction model is that with interpretability increased, model performance will be inevitably compromised because opaque models and features are discarded in the very first place. We propose to alter the rating prediction problem slightly into a polarity prediction problem, which still provide strong enough supervision signal on a user’s overall attitude towards a business, while allowing the model to achieve reasonable performance given relatively simplistic and coarse features.

Our contribution in this work is manifolds:

1. We proposed a novel task - to interpret rating prediction models;
2. We proposed to use random forest as classification algorithm, and reduced the rating prediction to a binary classification problem to balance the trade-off between performance and interpretability;
3. We proposed to leverage graph-based features to analyze community’s impact on ratings;
4. We hand-crafted and fine-tuned simple yet effective text-based features that break down semantics of user reviews to meaningful aspects.

The remaining parts are organized as follows: Section II provides background on Yelp dataset, random forest, and relevant graph algorithms; Section III introduces the features we use and our method to interpret the model; Section IV describes our experiments in detail and provide in-depth analysis on experimental results; Section V refers readers to related works; Section VI concludes this work.

## II. BACKGROUND

### Yelp Dataset

Yelp Dataset is a dataset containing users’ reviews and ratings of businesses. The dataset models users, businesses, and reviews as entities and relations. *User* is an entity that represent a yelp user, describing its social network and interactions as captured by its activity on Yelp. *Business* is

an entity that represent a merchant, with attributes like location and overall ratings. *Review* is a relation that depicts the interaction between a *User* and a *Business*, recording the user's rating to the business, with review attached as text.

### **Neo4j**

Neo4j is a graph database that supports the ACID-compliant transaction. It's also a database with native graph storage and processing. We used Neo4j to store, process and get important graph features from yelp datasets.

### **Random Forests**

Random forests are a learning method based on decision trees. Decision tree is a transparent classification model, splitting data into subsets on criteria understandable by human being. One problem with decision trees is that when they grow to certain depth, they easily overfit the training data.

By ensembling multiple decision trees, random forests is effective in circumventing the problem of overfitting. Random forests are a way of averaging multiple deep decision trees, trained on different parts of the same training set, with the goal of reducing the variance. It uses a modified tree learning algorithm that selects, at each candidate split in the learning process, a random subset of the features. This comes at the expense of a small increase in the bias and some loss of interpretability, but generally greatly boosts the performance in the final model.

One variation of random forests being used in this work is *extremely randomized trees*, or ExtraTrees. While similar to ordinary random forests in that they are an ensemble of individual trees, there are two main differences: first, each tree is trained using the whole learning sample (rather than a bootstrap sample), and second, the top-down splitting in the tree learner is randomized.

### **Graph Algorithms**

#### **1. Louvain method of community detection**

It's an algorithm for communities detection by the way of maximizing the modularity of each community, proposed by Blondel, Vincent D., et al [1]. The basic idea is in two steps: Firstly assign the community of nodes based in a greedy way for local optimizations. Secondly, create a coarse-grained network based on step one. Repeat those two steps until good optimizations achieved.

#### **2. PageRank**

PageRank algorithm is proposed by Brin, Sergey, and Lawrence Page [2], which measures the connectivity and influence of nodes. It's basically an algorithm that analyzes the importance of a web page by considering the number

and quality of hyperlinks between web pages. The basic assumption is that more important pages are referenced more compared with other pages. Also, It can be treated as a probability transition problem of Markov chain. The output of the algorithm will be a probability distribution of the likelihood that a user randomly clicking on links that arrive at the particular page.

#### **3. Betweenness**

Betweenness centrality indicates how the node in the graph can be served as a bridge to connect different parts of the graph and how nodes stand between each other. Higher betweenness means more information flows over that node. It was proposed by Linton C. Freeman[3]. The basic idea is to find the shortest path of pairs of nodes in the graph, the node lies in more paths indicates higher betweenness.

#### **4. Closeness**

Closeness centrality indicates how the node is close to all other nodes. Higher closeness means information could be spread more efficiently. It's proposed by Alex Bavelas [4] and the closeness of a node is the reciprocal of the sum of the length of the shortest paths between this node and all other nodes.

## **III. APPROACH**

### **Text Features**

One crucial part of contextual information based on which a rating prediction model could work is review texts. There exists innumerable feature extraction methods for texts, one of the most effective being pre-trained word embeddings. Others include TF-IDF, syntactic tags, etc. Despite these methods being notably sophisticated, we propose to extract features from basic n-grams to ensure transparency and hence interpretability of the whole model.

In our approach, we define *catch-phrases* as those n-grams that 1. have frequent occurrence across all reviews, and 2. signify a user's attitude towards certain aspect about the restaurant. For one example, the unigram 'delicious' is obviously talking positively about the food. For another example, 'seated right away' is apparently a compliment to the restaurant's efficiency.

To ensure all catch-phrases are meaningful and covers most reviews in the dataset, we manually selected a collection of them from the top-ranking unigrams, bigrams and trigrams.

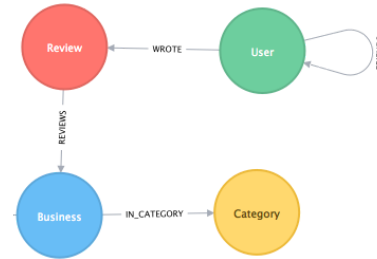
These catchphrases are then grouped based on their semantics, into a total of 13 categories. We then convert them into 13 features, with value 1 denoting positive, -1 denoting negative, and 0 denoting neutral in each corresponding category. The table below are some examples of catchphrases in each category.

Category	Positive Catch-phrases	Negative Catch-phrases
Food overall	delicious, authentic, amazing food	—
Special need	gluten free, dog friendly, open 24 hours	—
Alcohol	great beer selection	—
Food amount	big breakfast	—
Service overall	staff friendly, excellent customer service	worst service ever, rude
Service efficiency	seated right away, food came quickly	took long time, took forever
Will return or not	definitely come back	never go back
Will recommend or not	highly recommend	—
Restaurant overall	one favorite restaurants, worth wait	zero stars, worst experience
Location	within walking distance, right across street	little hard find
Place / Atmosphere	great atmosphere, feel right home	—
Price	worth every penny, reasonable prices	prices bit high
Sanitation	clean	—

### Graph Features

One problem we try to answer is how different communities react to different types of restaurant. To this end, we also incorporate graph features in our model.

The graph model we are using in this project is shown below: users could be friends of each other. A user could write a review that reviews a restaurant (business). A restaurant could be in multiple categories.



The features below are the graph features we processed and used in the project:

1. User Community: found by Louvain algorithm
2. User Centrality: PageRank
3. User Centrality: Betweenness
4. User Centrality: Closeness
5. Key opinion leader: top 500 PageRank users
6. Restaurant favored by different community (top 17 communities): Whether the average rating of this restaurant by users in this community, more than, user's total average rating in this community.
7. Key opinion leader contributes to the Review: Whether any key opinion leader followed by this user wrote reviews to this restaurant (before this user wrote the review to this restaurant).
8. Review's social rating: average social rating of 1st & 2nd degree social circle of this user who reviewed this restaurant (before this user wrote the review to this restaurant).

### Features come with datasets

There are many good features that already come with yelp datasets. We used the following features in the project:

1. User: received elite endorsement, average rating star
2. Restaurant: zip code, average star, category, city
3. Review: raw text (used for text features), review stars (converted to polarity)

### Interpretation Method

In order to interpret the prediction behavior of our model and understand what is the most important factor that

influences its decision, we calculate an importance score for each feature involved.

The method we used to compute importance is simple, yet effective and intuitive. In decision trees, the dataset is split at each node based on a specific feature that reduces the impurity measure most. The more such a feature contribute to impurity decrease, the more important it is considered to be. We compute a ‘mean impurity decrease’ as the total decrease in node impurity weighted by the probability of reaching that node, where that probability is approximated by the proportion of samples reaching that node averaged over all trees in the forest.

#### IV. EXPERIMENTS

##### Data Preprocessing

The original Yelp dataset is large and contains not only restaurant data. To make it easier for analysis, we restricted our experiment to a subset of business which are all restaurants from the city of Phoenix. In this way, we have 376,172 reviews, 142,286 users and 3,833 restaurants.

For the reason stated in Section I, we binarized the rating attached to each review in the following way: For 1, 2 and 3 stars, we consider them to be negative examples of reviews; For 4 and 5 stars, we consider them to be positive examples of reviews. This is because we discover that the mean rating is at about 4.5, which indicate that there is a bias toward high rating.

For review texts, we applied simple text normalization tricks such as lowercasing, tokenizing, and stopwords removal. We then extracted all unigram, bigram and trigrams separately and manually selected the most indicative catch-phrases.

##### Experimental Settings

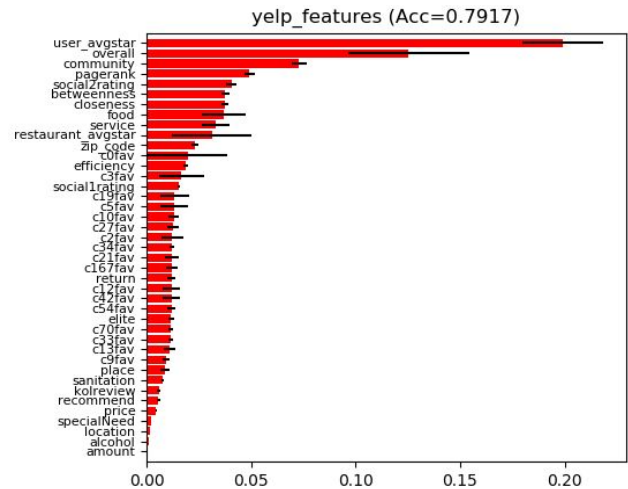
To dig into the impact of different factors have on ratings of users from different communities on different type (category) of restaurants, we performed multiple groups of experiments, each on a subset of reviews. These subsets are selected according to either the restaurant type or the identified user community or both. This results in 44 subsets. We also run our experiment on the full data, which add up to 45 group of experiments.

##### Results and Analysis

Most groups of experiments reached accuracies of around 80%, which is a fair enough performance to conclude that the classifier is learning something meaningful. We then calculated importance of each feature for all groups and had some interesting findings.

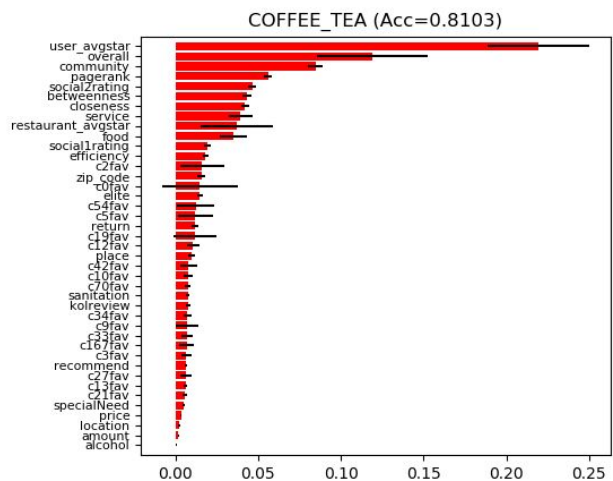
Here, we select several groups of experiments as example for analysis. The full experimental result can be found in the table in Appendix.

Ex. 1



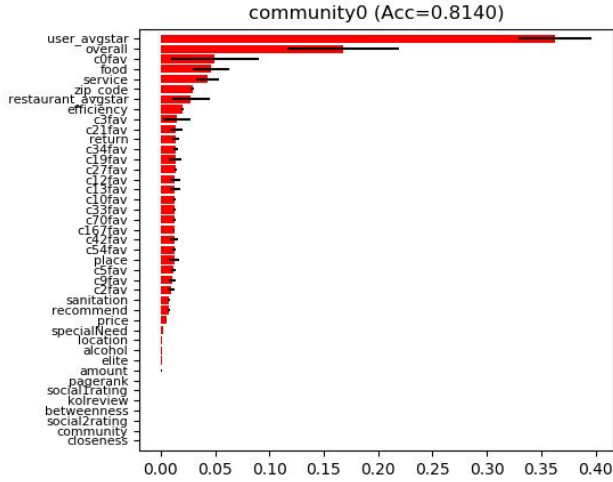
The Figure above displays the importance of features in the classifier trained on the full dataset. From the chart, we can see that ‘user\_avgstar’, which is the average rating that a user has given out historically, is attached the highest importance. Such phenomenon is recurring in a lot of groups of experiments. One possible interpretation to this could be that each user have a different baseline rating in his/her mind. For example, one person may give the restaurant he/she thinks is mediocre 3 stars, while another person thinks mediocre restaurants deserves 4 stars.

Ex. 2



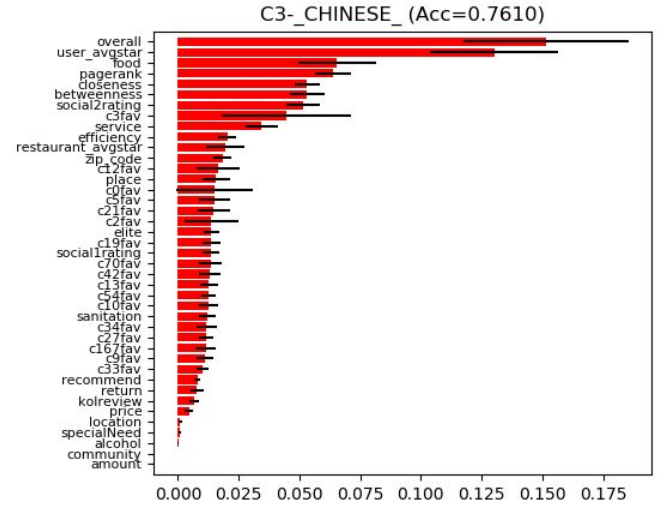
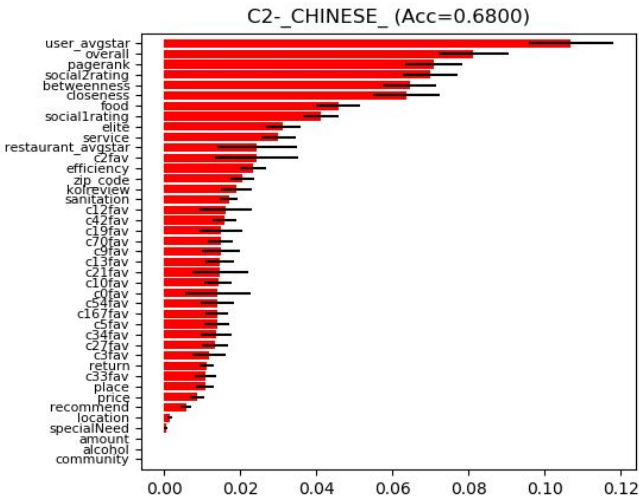
The Figure above displays the importance of features in the classifier trained on the subset containing coffee-tea restaurants. The chart shows that community-based features, such as community, pagerank, betweenness and closeness are playing important roles in classification. One explanation could be that coffee or tea consumption is sensitive to the community a user belongs to. This is reasonable because a lot of communities are featured by their cultural, ethnic, or religious identities.

Ex. 3



The Figure above displays the importance of features in the classifier trained on the subset containing users from community with index 0. Community-0 is a large community identified by our graph algorithm, with mixed users of different identities. Most possibly for this reason, restaurant ratings no longer depend heavily on a user's community background, but instead depend on some of those intrinsic qualities of that restaurant being evaluated. Therefore some of the most important features we see in the chart are 'overall' (restaurant overall), 'food' (food overall), 'service' (service overall), and 'efficiency' (service efficiency). These qualities are common judging standards that apply regardless of a user's cultural, ethnic, or religious identity.

Ex. 4



The two figures above displays the importance of features in the classifiers trained on the subset containing reviews on Chinese restaurants from users in community-2 and community-3. From the two charts, it is clear that 'kolreview', which is the key-opinion-leader's review, is having a drastically different degree of importance in the two communities. Such information could be considered helpful for restaurant owners to apply different strategies to the two different communities.

The results of these four experiments show some interesting facts. Since we have 45 experiments, we cannot show all of them here, you can find the full experimental result in the table in Appendix B.

## V. RELATED WORKS

Previously, there were some interesting works on review-based rating predictions [6] [7]. These works all focus mainly on the overall performance, without a good interpretability for their models.

Random forests were first proposed by [5], which is effective in solving the overfit problem for decision-tree based learning.

## VI. CONCLUSION

In this work, we build interpretable classifiers that achieves around 80% accuracy on Yelp dataset. We discover that for different communities or types of restaurant, factors contribute different weights to final rating. We also found that different users have different baseline rating scores. Restaurant owners can use our model to develop the customer strategy accommodating to different communities.

## REFERENCES

- [1] Blondel, Vincent D., et al. "Fast unfolding of communities in large networks." *Journal of statistical mechanics: theory and experiment* 2008.10 (2008): P10008.
- [2] Brin, Sergey, and Lawrence Page. "The anatomy of a large-scale hypertextual web search engine." *Computer networks and ISDN systems* 30.1-7 (1998): 107-117.
- [3] Freeman, Linton C. "A set of measures of centrality based on betweenness." *Sociometry* (1977): 35-41.
- [4] Alex Bavelas. Communication patterns in task-oriented groups. *J. Acoust. Soc. Am*, 22(6):725-730, 1950.
- [5] Ho, Tin Kam (1995). *Random Decision Forests*. Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, 14-16 August 1995. pp. 278-282.
- [6] Hadad, T. (2016). Review Based Rating Prediction. *CoRR*, *abs/1607.00024*.
- [7] Asghar, N. (2016). Yelp Dataset Challenge: Review Rating Prediction. *CoRR*, *abs/1605.05362*.

## APPENDIX A. SYSTEM MANUAL

Our experiment can be reproduced with the following steps:

1. Run 'filter\_jsons.py' to select only those reviews involving restaurants in Phoenix;
2. Run 'json2txt.py' to do review text preprocessing and convert data into text format;
3. Run 'extract\_text\_features.py' to extract text features and convert data into csv format;
4. Run scripts in 'graph\_scripts' to produce graph features;
5. Run 'add\_graph\_features.py' to combine text features and graph features;

6. Run 'run.py' to run all groups of experiments.

## APPENDIX B. FULL EXPERIMENTAL RESULTS

The table below contains results of all 45 groups of experiments. Each row contains result of one group of experiment; Each column contains statistics on one feature, including the feature's importance ranking, importance percentage score, and standard deviation of the importance score.

As is mentioned in Section IV, each group of experiment is run on one subset of data. The subset name follow these conventions: 'community#' denotes an identified community with unique index #; A name like 'FASTFOOD', 'CHINESE' denotes a certain type of restaurant; 'C#-TYPE\_NAME' denotes a subset that contains community#'s review on restaurants of type TYPE\_NAME.

Features follow these naming conventions: 'c#fav' means community#'s favoritism over this subset of restaurants; 'social#rating' means the user's #-degree connections' average rating; 'kolreview' means the key-opinion-leader has reviewed before this user and they are friends.

## APPENDIX C. CONTRIBUTIONS

Zhuoran Liu: Text features, experiments and analysis

Mingye Chen: Graph features, experiments and analysis

**Please check the Appendix B in the next page**

Subset	Accuracy	c0fav	social2rating	user_avgstar	closeness	community	c10fav	c5fav	c167fav
Full Data	0.7917	12 (0.0198±0.0190)	5 (0.0406±0.0023)	1 (0.1991±0.0194)	7 (0.0376±0.0018)	3 (0.0730±0.0036)	18 (0.0130±0.0023)	17 (0.0132±0.0066)	23 (0.0121±0.0028)
community0	0.814	3 (0.0499±0.0404)	39 (0.0000±0.0000)	1 (0.3628±0.0335)	41 (0.0000±0.0000)	40 (0.0000±0.0000)	17 (0.0128±0.0012)	24 (0.0117±0.0024)	20 (0.0123±0.0009)
community13	0.8135	29 (0.0139±0.0107)	5 (0.0396±0.0027)	1 (0.1233±0.0256)	6 (0.0361±0.0032)	41 (0.0000±0.0000)	12 (0.0201±0.0021)	20 (0.0180±0.0033)	22 (0.0177±0.0027)
community19	0.7866	12 (0.0221±0.0171)	6 (0.0380±0.0044)	2 (0.1003±0.0160)	8 (0.0296±0.0046)	41 (0.0000±0.0000)	14 (0.0211±0.0026)	24 (0.0163±0.0051)	15 (0.0203±0.0051)
community2	0.7285	28 (0.0157±0.0111)	3 (0.0709±0.0042)	1 (0.1091±0.0051)	6 (0.0600±0.0025)	41 (0.0000±0.0000)	15 (0.0173±0.0013)	29 (0.0149±0.0014)	26 (0.0160±0.0011)
community3	0.8025	13 (0.0184±0.0169)	5 (0.0475±0.0030)	1 (0.1553±0.0205)	6 (0.0474±0.0030)	41 (0.0000±0.0000)	17 (0.0163±0.0011)	28 (0.0139±0.0014)	25 (0.0153±0.0009)
community42	0.8059	24 (0.0156±0.0143)	9 (0.0299±0.0024)	2 (0.1292±0.0313)	10 (0.0297±0.0025)	41 (0.0000±0.0000)	17 (0.0179±0.0018)	22 (0.0160±0.0013)	19 (0.0174±0.0021)
community5	0.8086	12 (0.0211±0.0193)	5 (0.0429±0.0040)	3 (0.1025±0.0156)	7 (0.0349±0.0041)	41 (0.0000±0.0000)	19 (0.0194±0.0026)	1 (0.1120±0.0346)	21 (0.0191±0.0030)
community9	0.8356	12 (0.0215±0.0245)	9 (0.0337±0.0024)	1 (0.1487±0.0292)	5 (0.0375±0.0023)	41 (0.0000±0.0000)	16 (0.0178±0.0018)	15 (0.0180±0.0055)	25 (0.0158±0.0022)
FASTFOOD	0.7781	11 (0.0294±0.0344)	10 (0.0328±0.0050)	1 (0.1574±0.0323)	9 (0.0334±0.0024)	4 (0.0544±0.0032)	23 (0.0116±0.0046)	18 (0.0155±0.0086)	33 (0.0071±0.0010)
ITALIAN	0.7934	16 (0.0140±0.0127)	5 (0.0486±0.0021)	1 (0.2294±0.0189)	7 (0.0437±0.0021)	3 (0.0876±0.0035)	24 (0.0081±0.0017)	23 (0.0084±0.0024)	25 (0.0079±0.0010)
JAPANESE	0.7578	11 (0.0198±0.0201)	5 (0.0496±0.0037)	1 (0.1995±0.0214)	6 (0.0473±0.0040)	3 (0.0852±0.0047)	31 (0.0062±0.0012)	33 (0.0058±0.0019)	27 (0.0076±0.0027)
MEXICAN	0.7881	11 (0.0235±0.0242)	5 (0.0431±0.0033)	1 (0.2091±0.0182)	8 (0.0407±0.0019)	3 (0.0825±0.0026)	26 (0.0097±0.0027)	23 (0.0098±0.0031)	33 (0.0085±0.0010)
SALAD	0.7782	13 (0.0169±0.0175)	5 (0.0494±0.0030)	1 (0.2154±0.0173)	7 (0.0451±0.0030)	3 (0.0875±0.0038)	32 (0.0068±0.0011)	33 (0.0067±0.0035)	21 (0.0084±0.0015)
AMERICAN-TRADITIONAL	0.7743	12 (0.0187±0.0236)	5 (0.0436±0.0020)	1 (0.2060±0.0257)	8 (0.0396±0.0020)	3 (0.0800±0.0033)	24 (0.0099±0.0049)	28 (0.0094±0.0017)	33 (0.0080±0.0009)
CHINESE	0.7413	12 (0.0183±0.0209)	6 (0.0422±0.0024)	1 (0.1897±0.0207)	7 (0.0403±0.0023)	3 (0.0796±0.0037)	28 (0.0088±0.0033)	30 (0.0088±0.0021)	27 (0.0091±0.0027)
COFFEE_TEA	0.8103	15 (0.0148±0.0228)	5 (0.0465±0.0024)	1 (0.2195±0.0307)	7 (0.0421±0.0024)	3 (0.0847±0.0044)	24 (0.0079±0.0028)	18 (0.0119±0.0106)	31 (0.0068±0.0043)
C0-_AMERICANTRADITIONAL_	0.8025	4 (0.0477±0.0456)	37 (0.0000±0.0000)	1 (0.3780±0.0315)	39 (0.0000±0.0000)	38 (0.0000±0.0000)	22 (0.0096±0.0017)	13 (0.0119±0.0053)	17 (0.0108±0.0072)
C0-_CHINESE_	0.818	5 (0.0412±0.0374)	36 (0.0000±0.0000)	1 (0.3450±0.0337)	38 (0.0000±0.0000)	37 (0.0000±0.0000)	24 (0.0101±0.0022)	22 (0.0104±0.0021)	19 (0.0114±0.0059)
C0-_COFFEE_TEA_	0.8109	3 (0.0569±0.0497)	37 (0.0000±0.0000)	1 (0.4037±0.0402)	39 (0.0000±0.0000)	38 (0.0000±0.0000)	18 (0.0096±0.0032)	15 (0.0106±0.0098)	28 (0.0061±0.0020)
C0-_FASTFOOD_	0.831	3 (0.0889±0.0629)	39 (0.0000±0.0000)	1 (0.2690±0.0568)	35 (0.0000±0.0000)	40 (0.0000±0.0000)	16 (0.0141±0.0074)	10 (0.0170±0.0083)	22 (0.0099±0.0060)
C0-_ITALIAN_	0.7847	5 (0.0334±0.0272)	40 (0.0000±0.0000)	1 (0.4295±0.0292)	37 (0.0000±0.0000)	36 (0.0000±0.0000)	19 (0.0093±0.0014)	16 (0.0104±0.0048)	23 (0.0086±0.0021)
C0-_JAPANESE_	0.7791	5 (0.0403±0.0444)	34 (0.0000±0.0000)	1 (0.3971±0.0338)	38 (0.0000±0.0000)	37 (0.0000±0.0000)	24 (0.0083±0.0040)	19 (0.0090±0.0061)	25 (0.0083±0.0026)
C0-_MEXICAN_	0.8131	4 (0.0475±0.0371)	36 (0.0000±0.0000)	1 (0.3820±0.0348)	40 (0.0000±0.0000)	39 (0.0000±0.0000)	14 (0.0121±0.0030)	17 (0.0115±0.0034)	25 (0.0100±0.0014)
C0-_SALAD_	0.764	5 (0.0383±0.0307)	39 (0.0000±0.0000)	1 (0.4137±0.0317)	36 (0.0000±0.0000)	40 (0.0000±0.0000)	12 (0.0113±0.0035)	25 (0.0091±0.0046)	21 (0.0096±0.0020)
C2-_AMERICANTRADITIONAL_	0.7073	14 (0.0209±0.0155)	3 (0.0732±0.0041)	1 (0.1136±0.0057)	5 (0.0649±0.0031)	41 (0.0000±0.0000)	20 (0.0139±0.0034)	23 (0.0135±0.0026)	32 (0.0122±0.0015)
C2-_CHINESE_	0.68	25 (0.0142±0.0087)	4 (0.0700±0.0071)	1 (0.1069±0.0112)	6 (0.0637±0.0087)	41 (0.0000±0.0000)	24 (0.0142±0.0037)	28 (0.0139±0.0031)	27 (0.0140±0.0030)
C2-_COFFEE_TEA_	0.7681	23 (0.0118±0.0120)	2 (0.0861±0.0085)	1 (0.1195±0.0110)	6 (0.0702±0.0053)	41 (0.0000±0.0000)	20 (0.0129±0.0030)	30 (0.0104±0.0030)	31 (0.0097±0.0028)
C2-_FASTFOOD_	0.7247	17 (0.0196±0.0136)	4 (0.0549±0.0037)	2 (0.0895±0.0123)	6 (0.0499±0.0058)	41 (0.0000±0.0000)	24 (0.0162±0.0022)	19 (0.0189±0.0041)	32 (0.0118±0.0028)
C2-_ITALIAN_	0.7692	20 (0.0134±0.0085)	2 (0.0828±0.0059)	1 (0.1265±0.0079)	6 (0.0696±0.0049)	41 (0.0000±0.0000)	24 (0.0125±0.0028)	28 (0.0109±0.0027)	21 (0.0133±0.0026)
C2-_JAPANESE_	0.6968	24 (0.0120±0.0111)	4 (0.0755±0.0069)	1 (0.1126±0.0076)	6 (0.0688±0.0076)	41 (0.0000±0.0000)	26 (0.0116±0.0025)	32 (0.0093±0.0030)	22 (0.0125±0.0039)
C2-_MEXICAN_	0.7284	18 (0.0160±0.0138)	2 (0.0732±0.0046)	1 (0.1123±0.0043)	5 (0.0639±0.0045)	41 (0.0000±0.0000)	22 (0.0143±0.0015)	28 (0.0129±0.0029)	26 (0.0138±0.0021)
C2-_SALAD_	0.7385	23 (0.0118±0.0087)	2 (0.0853±0.0060)	1 (0.1265±0.0091)	6 (0.0698±0.0055)	41 (0.0000±0.0000)	27 (0.0114±0.0023)	32 (0.0084±0.0022)	18 (0.0141±0.0028)
C3-_AMERICANTRADITIONAL_	0.7641	13 (0.0172±0.0201)	5 (0.0507±0.0034)	1 (0.1636±0.0231)	4 (0.0519±0.0043)	41 (0.0000±0.0000)	23 (0.0133±0.0054)	29 (0.0118±0.0022)	30 (0.0118±0.0039)
C3-_CHINESE_	0.761	15 (0.0151±0.0158)	7 (0.0514±0.0069)	2 (0.1302±0.0265)	5 (0.0532±0.0053)	40 (0.0000±0.0000)	26 (0.0126±0.0039)	16 (0.0151±0.0064)	30 (0.0114±0.0040)
C3-_COFFEE_TEA_	0.8042	12 (0.0212±0.0260)	4 (0.0623±0.0051)	1 (0.1628±0.0252)	6 (0.0584±0.0061)	41 (0.0000±0.0000)	20 (0.0127±0.0047)	27 (0.0104±0.0076)	31 (0.0087±0.0029)
C3-_FASTFOOD_	0.8045	11 (0.0290±0.0277)	10 (0.0324±0.0065)	2 (0.1096±0.0358)	9 (0.0363±0.0043)	40 (0.0000±0.0000)	24 (0.0126±0.0035)	16 (0.0177±0.0077)	30 (0.0103±0.0032)
C3-_ITALIAN_	0.7937	22 (0.0123±0.0095)	4 (0.0570±0.0044)	1 (0.1714±0.0241)	5 (0.0565±0.0044)	41 (0.0000±0.0000)	18 (0.0132±0.0020)	27 (0.0117±0.0025)	21 (0.0125±0.0021)
C3-_JAPANESE_	0.7972	14 (0.0165±0.0181)	6 (0.0571±0.0062)	1 (0.1711±0.0190)	4 (0.0609±0.0071)	41 (0.0000±0.0000)	25 (0.0104±0.0033)	35 (0.0079±0.0025)	17 (0.0134±0.0047)
C3-_MEXICAN_	0.8025	14 (0.0156±0.0190)	6 (0.0532±0.0042)	1 (0.1598±0.0218)	4 (0.0552±0.0037)	41 (0.0000±0.0000)	18 (0.0142±0.0034)	27 (0.0124±0.0017)	24 (0.0126±0.0017)
C3-_SALAD_	0.7993	12 (0.0194±0.0155)	4 (0.0631±0.0062)	1 (0.1602±0.0194)	6 (0.0580±0.0059)	41 (0.0000±0.0000)	26 (0.0103±0.0023)	33 (0.0083±0.0021)	18 (0.0132±0.0030)
C42-_MEXICAN_	0.8033	12 (0.0197±0.0210)	9 (0.0357±0.0049)	2 (0.1541±0.0392)	5 (0.0401±0.0056)	40 (0.0000±0.0000)	15 (0.0164±0.0047)	28 (0.0129±0.0042)	24 (0.0139±0.0049)
C9-_AMERICANTRADITIONAL_	0.7987	10 (0.0249±0.0345)	9 (0.0384±0.0065)	2 (0.1374±0.0366)	7 (0.0410±0.0058)	41 (0.0000±0.0000)	22 (0.0156±0.0070)	24 (0.0148±0.0039)	23 (0.0152±0.0038)
C9-_ITALIAN_	0.8178	16 (0.0170±0.0166)	6 (0.0432±0.0056)	2 (0.1459±0.0387)	9 (0.0394±0.0079)	41 (0.0000±0.0000)	14 (0.0173±0.0060)	15 (0.0172±0.0117)	18 (0.0161±0.0074)
C9-_MEXICAN_	0.8033	12 (0.0197±0.0210)	9 (0.0357±0.0049)	2 (0.1541±0.0392)	5 (0.0401±0.0056)	40 (0.0000±0.0000)	15 (0.0164±0.0047)	28 (0.0129±0.0042)	24 (0.0139±0.0049)



Subset	Accuracy	c34fav	c12fav	alcohol	service	elite	c21fav	betweenness	location
Full Data	0.7917	21 (0.0122±0.0011)	25 (0.0119±0.0040)	40 (0.0011±0.0001)	9 (0.0331±0.0066)	28 (0.0118±0.0013)	22 (0.0121±0.0035)	6 (0.0376±0.0019)	39 (0.0014±0.0001)
community0	0.814	12 (0.0136±0.0019)	15 (0.0131±0.0044)	32 (0.0011±0.0001)	5 (0.0432±0.0102)	33 (0.0009±0.0001)	10 (0.0141±0.0059)	38 (0.0000±0.0000)	31 (0.0012±0.0001)
community13	0.8135	16 (0.0191±0.0019)	23 (0.0175±0.0032)	38 (0.0011±0.0003)	9 (0.0315±0.0064)	32 (0.0104±0.0022)	15 (0.0192±0.0026)	7 (0.0357±0.0035)	40 (0.0005±0.0003)
community19	0.7866	13 (0.0212±0.0032)	17 (0.0199±0.0047)	38 (0.0012±0.0005)	10 (0.0270±0.0054)	22 (0.0170±0.0035)	18 (0.0198±0.0056)	9 (0.0283±0.0046)	40 (0.0006±0.0004)
community2	0.7285	21 (0.0167±0.0015)	24 (0.0163±0.0029)	39 (0.0014±0.0001)	12 (0.0256±0.0030)	13 (0.0234±0.0022)	23 (0.0164±0.0031)	5 (0.0603±0.0028)	38 (0.0021±0.0002)
community3	0.8025	18 (0.0162±0.0014)	22 (0.0155±0.0032)	39 (0.0013±0.0002)	9 (0.0341±0.0051)	31 (0.0116±0.0012)	19 (0.0159±0.0025)	4 (0.0478±0.0032)	38 (0.0017±0.0002)
community42	0.8059	13 (0.0185±0.0021)	18 (0.0175±0.0046)	38 (0.0010±0.0004)	6 (0.0313±0.0093)	30 (0.0111±0.0016)	20 (0.0173±0.0037)	7 (0.0302±0.0039)	39 (0.0008±0.0004)
community5	0.8086	13 (0.0204±0.0033)	18 (0.0194±0.0028)	38 (0.0018±0.0007)	10 (0.0310±0.0058)	27 (0.0148±0.0019)	20 (0.0192±0.0050)	8 (0.0337±0.0035)	39 (0.0016±0.0006)
community9	0.8356	17 (0.0170±0.0015)	21 (0.0164±0.0053)	39 (0.0006±0.0002)	6 (0.0370±0.0129)	28 (0.0128±0.0018)	19 (0.0168±0.0047)	8 (0.0363±0.0034)	38 (0.0007±0.0002)
FASTFOOD	0.7781	26 (0.0102±0.0011)	30 (0.0075±0.0015)	41 (0.0001±0.0001)	6 (0.0482±0.0105)	20 (0.0151±0.0025)	25 (0.0106±0.0043)	8 (0.0334±0.0022)	39 (0.0011±0.0003)
ITALIAN	0.7934	28 (0.0075±0.0010)	19 (0.0093±0.0053)	40 (0.0006±0.0001)	8 (0.0390±0.0059)	12 (0.0157±0.0013)	27 (0.0076±0.0038)	6 (0.0438±0.0016)	39 (0.0018±0.0003)
JAPANESE	0.7578	34 (0.0055±0.0014)	29 (0.0072±0.0021)	40 (0.0006±0.0002)	9 (0.0392±0.0054)	13 (0.0184±0.0020)	20 (0.0106±0.0079)	7 (0.0462±0.0029)	38 (0.0021±0.0004)
MEXICAN	0.7881	24 (0.0097±0.0024)	30 (0.0089±0.0049)	40 (0.0007±0.0001)	9 (0.0324±0.0051)	15 (0.0144±0.0015)	31 (0.0088±0.0010)	7 (0.0408±0.0021)	39 (0.0012±0.0002)
SALAD	0.7782	24 (0.0079±0.0016)	31 (0.0069±0.0023)	41 (0.0009±0.0002)	9 (0.0374±0.0057)	14 (0.0164±0.0009)	23 (0.0081±0.0034)	6 (0.0452±0.0025)	39 (0.0014±0.0003)
AMERICAN-TRADITIONAL	0.7743	30 (0.0087±0.0009)	21 (0.0120±0.0087)	39 (0.0018±0.0002)	9 (0.0349±0.0034)	15 (0.0146±0.0018)	29 (0.0091±0.0033)	7 (0.0399±0.0017)	40 (0.0017±0.0001)
CHINESE	0.7413	25 (0.0096±0.0060)	19 (0.0113±0.0070)	40 (0.0002±0.0001)	9 (0.0328±0.0054)	14 (0.0152±0.0021)	22 (0.0109±0.0077)	8 (0.0395±0.0027)	38 (0.0009±0.0003)
COFFEE_TEA	0.8103	28 (0.0073±0.0023)	21 (0.0107±0.0037)	41 (0.0005±0.0001)	8 (0.0394±0.0074)	16 (0.0147±0.0017)	36 (0.0056±0.0012)	6 (0.0433±0.0026)	39 (0.0023±0.0005)
C0-_AMERICANTRADITIONAL_	0.8025	21 (0.0103±0.0019)	14 (0.0119±0.0069)	31 (0.0023±0.0004)	5 (0.0458±0.0084)	33 (0.0009±0.0002)	23 (0.0095±0.0018)	40 (0.0000±0.0000)	32 (0.0018±0.0003)
C0-_CHINESE_	0.818	21 (0.0104±0.0065)	23 (0.0103±0.0038)	32 (0.0002±0.0002)	4 (0.0527±0.0102)	30 (0.0006±0.0003)	13 (0.0125±0.0096)	39 (0.0000±0.0000)	31 (0.0004±0.0003)
C0-_COFFEE_TEA_	0.8109	16 (0.0105±0.0059)	12 (0.0121±0.0048)	33 (0.0008±0.0004)	4 (0.0535±0.0149)	34 (0.0004±0.0002)	19 (0.0092±0.0038)	40 (0.0000±0.0000)	32 (0.0014±0.0005)
C0-_FASTFOOD_	0.831	20 (0.0115±0.0021)	24 (0.0091±0.0017)	34 (0.0000±0.0000)	5 (0.0671±0.0210)	32 (0.0009±0.0005)	17 (0.0139±0.0066)	36 (0.0000±0.0000)	31 (0.0009±0.0005)
C0-_ITALIAN_	0.7847	21 (0.0091±0.0018)	18 (0.0093±0.0037)	32 (0.0009±0.0002)	3 (0.0554±0.0122)	31 (0.0010±0.0003)	22 (0.0087±0.0030)	35 (0.0000±0.0000)	33 (0.0007±0.0002)
C0-_JAPANESE_	0.7791	26 (0.0079±0.0028)	20 (0.0090±0.0034)	33 (0.0000±0.0001)	3 (0.0570±0.0139)	32 (0.0004±0.0004)	14 (0.0125±0.0106)	39 (0.0000±0.0000)	30 (0.0026±0.0005)
C0-_MEXICAN_	0.8131	20 (0.0110±0.0015)	16 (0.0116±0.0053)	31 (0.0011±0.0002)	5 (0.0453±0.0071)	33 (0.0009±0.0003)	18 (0.0114±0.0038)	41 (0.0000±0.0000)	30 (0.0012±0.0003)
C0-_SALAD_	0.764	13 (0.0111±0.0029)	19 (0.0098±0.0035)	34 (0.0007±0.0003)	4 (0.0531±0.0103)	31 (0.0019±0.0004)	14 (0.0109±0.0051)	41 (0.0000±0.0000)	32 (0.0009±0.0003)
C2-_AMERICANTRADITIONAL_	0.7073	22 (0.0138±0.0016)	18 (0.0152±0.0060)	37 (0.0026±0.0005)	9 (0.0336±0.0034)	10 (0.0291±0.0039)	25 (0.0134±0.0025)	6 (0.0643±0.0039)	39 (0.0022±0.0006)
C2-_CHINESE_	0.68	29 (0.0137±0.0041)	17 (0.0162±0.0069)	40 (0.0000±0.0000)	10 (0.0301±0.0044)	9 (0.0314±0.0045)	23 (0.0148±0.0075)	5 (0.0645±0.0070)	37 (0.0016±0.0007)
C2-_COFFEE_TEA_	0.7681	29 (0.0105±0.0026)	17 (0.0151±0.0041)	40 (0.0004±0.0003)	8 (0.0371±0.0064)	11 (0.0283±0.0047)	24 (0.0111±0.0032)	5 (0.0705±0.0061)	38 (0.0031±0.0010)
C2-_FASTFOOD_	0.7247	21 (0.0170±0.0035)	31 (0.0128±0.0038)	40 (0.0000±0.0001)	10 (0.0322±0.0046)	14 (0.0269±0.0037)	23 (0.0167±0.0040)	5 (0.0523±0.0062)	38 (0.0014±0.0007)
C2-_ITALIAN_	0.7692	25 (0.0117±0.0020)	22 (0.0130±0.0032)	39 (0.0009±0.0004)	8 (0.0325±0.0052)	10 (0.0285±0.0044)	26 (0.0117±0.0044)	5 (0.0704±0.0059)	38 (0.0036±0.0009)
C2-_JAPANESE_	0.6968	29 (0.0104±0.0036)	25 (0.0120±0.0038)	40 (0.0002±0.0003)	9 (0.0373±0.0056)	10 (0.0307±0.0081)	19 (0.0134±0.0048)	5 (0.0739±0.0077)	37 (0.0023±0.0008)
C2-_MEXICAN_	0.7284	23 (0.0142±0.0019)	25 (0.0140±0.0019)	39 (0.0009±0.0003)	10 (0.0317±0.0041)	11 (0.0287±0.0031)	21 (0.0147±0.0024)	6 (0.0634±0.0033)	38 (0.0015±0.0004)
C2-_SALAD_	0.7385	17 (0.0144±0.0037)	29 (0.0107±0.0022)	40 (0.0012±0.0005)	9 (0.0338±0.0043)	10 (0.0257±0.0065)	24 (0.0117±0.0027)	5 (0.0705±0.0064)	38 (0.0021±0.0007)
C3-_AMERICANTRADITIONAL_	0.7641	21 (0.0135±0.0016)	15 (0.0164±0.0095)	39 (0.0014±0.0004)	9 (0.0413±0.0066)	22 (0.0134±0.0024)	28 (0.0122±0.0017)	6 (0.0499±0.0044)	38 (0.0017±0.0005)
C3-_CHINESE_	0.761	28 (0.0118±0.0042)	13 (0.0166±0.0089)	39 (0.0001±0.0001)	9 (0.0345±0.0067)	19 (0.0137±0.0032)	17 (0.0146±0.0066)	6 (0.0532±0.0073)	37 (0.0009±0.0008)
C3-_COFFEE_TEA_	0.8042	26 (0.0106±0.0047)	17 (0.0133±0.0029)	40 (0.0006±0.0005)	8 (0.0394±0.0068)	15 (0.0164±0.0027)	16 (0.0143±0.0066)	5 (0.0597±0.0072)	38 (0.0019±0.0009)
C3-_FASTFOOD_	0.8045	21 (0.0137±0.0042)	29 (0.0109±0.0036)	39 (0.0005±0.0004)	6 (0.0378±0.0128)	22 (0.0132±0.0032)	18 (0.0156±0.0029)	8 (0.0367±0.0056)	38 (0.0013±0.0007)
C3-_ITALIAN_	0.7937	26 (0.0118±0.0024)	19 (0.0129±0.0037)	39 (0.0007±0.0004)	7 (0.0402±0.0085)	16 (0.0137±0.0026)	28 (0.0116±0.0023)	6 (0.0564±0.0045)	38 (0.0020±0.0007)
C3-_JAPANESE_	0.7972	24 (0.0104±0.0029)	22 (0.0118±0.0040)	39 (0.0005±0.0005)	8 (0.0430±0.0068)	16 (0.0134±0.0033)	19 (0.0125±0.0065)	5 (0.0606±0.0088)	37 (0.0025±0.0011)
C3-_MEXICAN_	0.8025	23 (0.0128±0.0020)	28 (0.0124±0.0043)	39 (0.0007±0.0003)	9 (0.0363±0.0076)	31 (0.0118±0.0030)	25 (0.0125±0.0017)	5 (0.0545±0.0036)	38 (0.0009±0.0004)
C3-_SALAD_	0.7993	19 (0.0126±0.0036)	25 (0.0103±0.0024)	39 (0.0008±0.0005)	7 (0.0417±0.0052)	15 (0.0153±0.0025)	21 (0.0120±0.0034)	5 (0.0596±0.0036)	38 (0.0021±0.0006)
C42-_MEXICAN_	0.8033	17 (0.0163±0.0043)	25 (0.0139±0.0046)	39 (0.0001±0.0001)	8 (0.0376±0.0137)	23 (0.0141±0.0032)	22 (0.0141±0.0053)	7 (0.0393±0.0047)	37 (0.0010±0.0007)
C9-_AMERICANTRADITIONAL_	0.7987	25 (0.0143±0.0038)	20 (0.0162±0.0107)	38 (0.0008±0.0006)	5 (0.0440±0.0147)	26 (0.0141±0.0037)	18 (0.0171±0.0085)	8 (0.0410±0.0068)	37 (0.0010±0.0009)
C9-_ITALIAN_	0.8178	22 (0.0148±0.0045)	26 (0.0139±0.0059)	38 (0.0004±0.0005)	5 (0.0489±0.0119)	24 (0.0142±0.0039)	25 (0.0141±0.0048)	7 (0.0413±0.0086)	39 (0.0002±0.0003)
C9-_MEXICAN_	0.8033	17 (0.0163±0.0043)	25 (0.0139±0.0046)	39 (0.0001±0.0001)	8 (0.0376±0.0137)	23 (0.0141±0.0032)	22 (0.0141±0.0053)	7 (0.0393±0.0047)	37 (0.0010±0.0007)



Subset	Accuracy	recommend	c3fav	price	specialNeed	zip_code	return	c33fav	food
Full Data	0.7917	36 (0.0057±0.0007)	14 (0.0168±0.0110)	37 (0.0046±0.0002)	38 (0.0021±0.0001)	11 (0.0231±0.0018)	24 (0.0120±0.0019)	30 (0.0114±0.0011)	8 (0.0370±0.0104)
community0	0.814	28 (0.0074±0.0007)	9 (0.0149±0.0129)	29 (0.0050±0.0007)	30 (0.0021±0.0002)	6 (0.0292±0.0015)	11 (0.0141±0.0030)	18 (0.0125±0.0008)	4 (0.0463±0.0167)
community13	0.8135	34 (0.0064±0.0011)	28 (0.0140±0.0059)	36 (0.0043±0.0007)	37 (0.0021±0.0004)	10 (0.0305±0.0019)	30 (0.0118±0.0026)	18 (0.0185±0.0021)	8 (0.0342±0.0097)
community19	0.7866	33 (0.0075±0.0019)	31 (0.0115±0.0039)	36 (0.0038±0.0011)	37 (0.0018±0.0007)	7 (0.0336±0.0036)	27 (0.0137±0.0031)	19 (0.0192±0.0027)	4 (0.0452±0.0067)
community2	0.7285	35 (0.0061±0.0009)	32 (0.0106±0.0020)	36 (0.0053±0.0003)	37 (0.0024±0.0003)	10 (0.0309±0.0015)	31 (0.0108±0.0009)	25 (0.0162±0.0014)	9 (0.0317±0.0055)
community3	0.8025	34 (0.0058±0.0010)	8 (0.0358±0.0194)	36 (0.0043±0.0004)	37 (0.0022±0.0002)	10 (0.0287±0.0016)	30 (0.0119±0.0021)	23 (0.0155±0.0012)	7 (0.0376±0.0085)
community42	0.8059	34 (0.0062±0.0010)	25 (0.0155±0.0106)	35 (0.0029±0.0007)	37 (0.0010±0.0004)	8 (0.0300±0.0032)	31 (0.0110±0.0023)	23 (0.0158±0.0023)	5 (0.0404±0.0141)
community5	0.8086	33 (0.0075±0.0016)	29 (0.0140±0.0059)	36 (0.0053±0.0012)	37 (0.0026±0.0009)	9 (0.0310±0.0033)	30 (0.0128±0.0029)	16 (0.0202±0.0035)	6 (0.0354±0.0070)
community9	0.8356	33 (0.0070±0.0010)	30 (0.0120±0.0059)	36 (0.0027±0.0004)	37 (0.0020±0.0005)	10 (0.0276±0.0020)	31 (0.0109±0.0017)	24 (0.0162±0.0017)	7 (0.0369±0.0143)
FASTFOOD	0.7781	36 (0.0047±0.0006)	17 (0.0160±0.0114)	37 (0.0033±0.0005)	38 (0.0015±0.0003)	13 (0.0189±0.0018)	31 (0.0075±0.0009)	32 (0.0075±0.0014)	3 (0.0688±0.0297)
ITALIAN	0.7934	30 (0.0073±0.0008)	17 (0.0111±0.0104)	37 (0.0052±0.0006)	38 (0.0031±0.0004)	15 (0.0141±0.0012)	14 (0.0143±0.0017)	31 (0.0073±0.0010)	9 (0.0330±0.0053)
JAPANESE	0.7578	22 (0.0097±0.0008)	21 (0.0099±0.0136)	28 (0.0075±0.0010)	39 (0.0017±0.0003)	18 (0.0117±0.0016)	14 (0.0166±0.0033)	32 (0.0060±0.0023)	8 (0.0413±0.0072)
MEXICAN	0.7881	35 (0.0072±0.0011)	19 (0.0112±0.0103)	37 (0.0047±0.0005)	38 (0.0013±0.0001)	14 (0.0158±0.0012)	16 (0.0140±0.0015)	34 (0.0074±0.0009)	6 (0.0424±0.0096)
SALAD	0.7782	29 (0.0073±0.0011)	26 (0.0077±0.0032)	37 (0.0054±0.0005)	38 (0.0031±0.0006)	16 (0.0131±0.0020)	15 (0.0141±0.0010)	35 (0.0062±0.0015)	8 (0.0415±0.0064)
AMERICAN-TRADITIONAL	0.7743	35 (0.0067±0.0008)	14 (0.0167±0.0172)	37 (0.0048±0.0003)	38 (0.0021±0.0002)	18 (0.0133±0.0012)	17 (0.0134±0.0021)	32 (0.0082±0.0009)	6 (0.0428±0.0121)
CHINESE	0.7413	35 (0.0080±0.0010)	20 (0.0112±0.0050)	37 (0.0064±0.0007)	39 (0.0007±0.0003)	15 (0.0126±0.0018)	17 (0.0117±0.0020)	33 (0.0085±0.0058)	4 (0.0681±0.0153)
COFFEE_TEA	0.8103	33 (0.0063±0.0009)	32 (0.0066±0.0033)	38 (0.0034±0.0005)	37 (0.0050±0.0007)	14 (0.0157±0.0022)	19 (0.0118±0.0018)	30 (0.0070±0.0033)	10 (0.0353±0.0083)
C0-_AMERICANTRADITIONAL_	0.8025	26 (0.0085±0.0013)	19 (0.0107±0.0126)	29 (0.0051±0.0005)	30 (0.0026±0.0005)	10 (0.0153±0.0021)	8 (0.0174±0.0035)	20 (0.0103±0.0020)	3 (0.0576±0.0182)
C0-_CHINESE_	0.818	17 (0.0116±0.0019)	8 (0.0156±0.0202)	27 (0.0096±0.0021)	33 (0.0002±0.0002)	12 (0.0140±0.0033)	11 (0.0144±0.0030)	28 (0.0082±0.0025)	3 (0.0780±0.0274)
C0-_COFFEE_TEA_	0.8109	23 (0.0083±0.0015)	13 (0.0115±0.0166)	30 (0.0033±0.0009)	29 (0.0057±0.0012)	8 (0.0194±0.0035)	10 (0.0132±0.0021)	24 (0.0083±0.0052)	5 (0.0470±0.0152)
C0-_FASTFOOD_	0.831	28 (0.0058±0.0011)	14 (0.0144±0.0092)	29 (0.0033±0.0010)	30 (0.0011±0.0007)	8 (0.0232±0.0029)	26 (0.0077±0.0017)	25 (0.0088±0.0017)	4 (0.0774±0.0393)
C0-_ITALIAN_	0.7847	15 (0.0110±0.0022)	11 (0.0134±0.0126)	29 (0.0054±0.0008)	30 (0.0029±0.0007)	9 (0.0168±0.0033)	8 (0.0196±0.0032)	17 (0.0094±0.0019)	4 (0.0460±0.0088)
C0-_JAPANESE_	0.7791	12 (0.0126±0.0030)	8 (0.0178±0.0200)	27 (0.0072±0.0014)	31 (0.0020±0.0006)	10 (0.0146±0.0049)	9 (0.0175±0.0044)	21 (0.0090±0.0065)	4 (0.0498±0.0083)
C0-_MEXICAN_	0.8131	28 (0.0088±0.0016)	10 (0.0142±0.0146)	29 (0.0053±0.0010)	32 (0.0010±0.0002)	8 (0.0196±0.0018)	9 (0.0162±0.0034)	27 (0.0095±0.0014)	3 (0.0489±0.0151)
C0-_SALAD_	0.764	23 (0.0095±0.0017)	11 (0.0134±0.0124)	29 (0.0057±0.0014)	30 (0.0027±0.0008)	9 (0.0156±0.0025)	8 (0.0156±0.0039)	26 (0.0087±0.0013)	3 (0.0537±0.0120)
C2-_AMERICANTRADITIONAL_	0.7073	35 (0.0077±0.0012)	34 (0.0099±0.0041)	36 (0.0058±0.0007)	38 (0.0024±0.0005)	15 (0.0205±0.0031)	29 (0.0129±0.0013)	27 (0.0132±0.0020)	8 (0.0364±0.0071)
C2-_CHINESE_	0.68	36 (0.0058±0.0014)	31 (0.0118±0.0043)	35 (0.0086±0.0018)	38 (0.0005±0.0005)	14 (0.0206±0.0030)	32 (0.0111±0.0019)	33 (0.0110±0.0028)	7 (0.0458±0.0057)
C2-_COFFEE_TEA_	0.7681	34 (0.0069±0.0016)	33 (0.0078±0.0025)	37 (0.0035±0.0011)	36 (0.0058±0.0015)	12 (0.0216±0.0029)	21 (0.0124±0.0018)	26 (0.0110±0.0048)	10 (0.0308±0.0055)
C2-_FASTFOOD_	0.7247	35 (0.0052±0.0009)	28 (0.0148±0.0029)	36 (0.0042±0.0015)	37 (0.0028±0.0010)	12 (0.0285±0.0042)	33 (0.0073±0.0014)	30 (0.0129±0.0026)	8 (0.0456±0.0119)
C2-_ITALIAN_	0.7692	35 (0.0081±0.0013)	34 (0.0087±0.0024)	36 (0.0060±0.0009)	37 (0.0041±0.0007)	12 (0.0224±0.0025)	16 (0.0143±0.0017)	23 (0.0126±0.0019)	9 (0.0308±0.0034)
C2-_JAPANESE_	0.6968	30 (0.0104±0.0023)	36 (0.0078±0.0045)	33 (0.0089±0.0022)	38 (0.0015±0.0007)	15 (0.0167±0.0029)	17 (0.0141±0.0028)	21 (0.0126±0.0049)	7 (0.0482±0.0072)
C2-_MEXICAN_	0.7284	35 (0.0083±0.0011)	33 (0.0104±0.0052)	36 (0.0055±0.0007)	37 (0.0021±0.0004)	12 (0.0231±0.0024)	31 (0.0112±0.0018)	29 (0.0121±0.0025)	9 (0.0337±0.0046)
C2-_SALAD_	0.7385	34 (0.0083±0.0022)	33 (0.0084±0.0025)	36 (0.0060±0.0014)	37 (0.0038±0.0009)	14 (0.0190±0.0033)	16 (0.0154±0.0019)	30 (0.0100±0.0024)	8 (0.0349±0.0040)
C3-_AMERICANTRADITIONAL_	0.7641	35 (0.0066±0.0012)	7 (0.0420±0.0307)	36 (0.0043±0.0009)	37 (0.0020±0.0004)	12 (0.0188±0.0020)	19 (0.0140±0.0035)	24 (0.0129±0.0019)	8 (0.0414±0.0096)
C3-_CHINESE_	0.761	33 (0.0081±0.0013)	8 (0.0447±0.0267)	36 (0.0045±0.0017)	38 (0.0007±0.0004)	12 (0.0183±0.0037)	34 (0.0078±0.0027)	32 (0.0102±0.0027)	3 (0.0655±0.0161)
C3-_COFFEE_TEA_	0.8042	32 (0.0083±0.0022)	7 (0.0417±0.0232)	36 (0.0058±0.0014)	37 (0.0040±0.0017)	11 (0.0216±0.0057)	28 (0.0093±0.0020)	23 (0.0118±0.0037)	9 (0.0328±0.0091)
C3-_FASTFOOD_	0.8045	35 (0.0045±0.0012)	4 (0.0709±0.0398)	34 (0.0046±0.0012)	37 (0.0017±0.0009)	13 (0.0198±0.0035)	33 (0.0068±0.0021)	27 (0.0113±0.0026)	3 (0.0856±0.0367)
C3-_ITALIAN_	0.7937	33 (0.0087±0.0024)	9 (0.0254±0.0139)	35 (0.0063±0.0014)	37 (0.0032±0.0006)	10 (0.0211±0.0027)	14 (0.0168±0.0037)	20 (0.0126±0.0021)	8 (0.0341±0.0069)
C3-_JAPANESE_	0.7972	33 (0.0082±0.0015)	9 (0.0321±0.0243)	28 (0.0095±0.0025)	38 (0.0018±0.0009)	11 (0.0200±0.0046)	26 (0.0104±0.0028)	20 (0.0125±0.0066)	7 (0.0482±0.0065)
C3-_MEXICAN_	0.8025	34 (0.0069±0.0013)	7 (0.0515±0.0288)	36 (0.0046±0.0008)	37 (0.0010±0.0004)	12 (0.0201±0.0029)	30 (0.0122±0.0013)	26 (0.0125±0.0021)	8 (0.0430±0.0081)
C3-_SALAD_	0.7993	35 (0.0069±0.0012)	9 (0.0303±0.0179)	36 (0.0047±0.0011)	37 (0.0029±0.0009)	13 (0.0183±0.0031)	16 (0.0147±0.0026)	29 (0.0094±0.0021)	8 (0.0390±0.0062)
C42-_MEXICAN_	0.8033	34 (0.0081±0.0021)	20 (0.0144±0.0146)	35 (0.0037±0.0020)	38 (0.0005±0.0005)	11 (0.0234±0.0042)	26 (0.0137±0.0032)	32 (0.0103±0.0026)	6 (0.0399±0.0127)
C9-_AMERICANTRADITIONAL_	0.7987	33 (0.0080±0.0021)	15 (0.0180±0.0175)	36 (0.0030±0.0015)	39 (0.0007±0.0006)	13 (0.0200±0.0043)	31 (0.0116±0.0024)	28 (0.0127±0.0033)	6 (0.0427±0.0184)
C9-_ITALIAN_	0.8178	32 (0.0095±0.0036)	17 (0.0167±0.0127)	37 (0.0026±0.0018)	35 (0.0045±0.0018)	11 (0.0227±0.0068)	33 (0.0083±0.0031)	20 (0.0158±0.0060)	8 (0.0400±0.0100)
C9-_MEXICAN_	0.8033	34 (0.0081±0.0021)	20 (0.0144±0.0146)	35 (0.0037±0.0020)	38 (0.0005±0.0005)	11 (0.0234±0.0042)	26 (0.0137±0.0032)	32 (0.0103±0.0026)	6 (0.0399±0.0127)

Subset	Accuracy	c19fav	c70fav	overall	efficiency	restaurant_avgstar	pagerank	c9fav	c13fav
Full Data	0.7917	16 (0.0135±0.0071)	29 (0.0117±0.0010)	2 (0.1254±0.0289)	13 (0.0187±0.0010)	10 (0.0313±0.0190)	4 (0.0493±0.0024)	32 (0.0093±0.0017)	31 (0.0112±0.0028)
community0	0.814	13 (0.0134±0.0060)	19 (0.0124±0.0013)	2 (0.1680±0.0516)	8 (0.0200±0.0014)	7 (0.0274±0.0173)	35 (0.0000±0.0000)	25 (0.0103±0.0033)	16 (0.0129±0.0045)
community13	0.8135	19 (0.0183±0.0026)	14 (0.0195±0.0028)	2 (0.1216±0.0405)	24 (0.0174±0.0020)	11 (0.0251±0.0075)	4 (0.0447±0.0045)	27 (0.0141±0.0026)	3 (0.1033±0.0421)
community19	0.7866	1 (0.1591±0.0515)	21 (0.0180±0.0026)	3 (0.0887±0.0248)	23 (0.0170±0.0036)	11 (0.0255±0.0086)	5 (0.0433±0.0034)	29 (0.0127±0.0031)	26 (0.0153±0.0028)
community2	0.7285	17 (0.0171±0.0051)	16 (0.0171±0.0011)	2 (0.0746±0.0098)	14 (0.0201±0.0018)	11 (0.0259±0.0080)	4 (0.0672±0.0023)	30 (0.0141±0.0033)	19 (0.0169±0.0051)
community3	0.8025	14 (0.0171±0.0048)	21 (0.0156±0.0014)	2 (0.1357±0.0271)	12 (0.0195±0.0012)	11 (0.0234±0.0036)	3 (0.0582±0.0035)	29 (0.0121±0.0019)	24 (0.0153±0.0057)
community42	0.8059	12 (0.0190±0.0059)	14 (0.0184±0.0018)	3 (0.1142±0.0406)	26 (0.0138±0.0021)	11 (0.0269±0.0069)	4 (0.0432±0.0038)	27 (0.0132±0.0027)	21 (0.0162±0.0023)
community5	0.8086	15 (0.0203±0.0073)	14 (0.0203±0.0032)	2 (0.1029±0.0271)	24 (0.0180±0.0024)	11 (0.0275±0.0152)	4 (0.0468±0.0035)	31 (0.0126±0.0022)	26 (0.0151±0.0017)
community9	0.8356	13 (0.0193±0.0087)	20 (0.0166±0.0013)	2 (0.1420±0.0425)	22 (0.0163±0.0015)	11 (0.0258±0.0103)	4 (0.0451±0.0035)	3 (0.0772±0.0348)	26 (0.0144±0.0029)
FASTFOOD	0.7781	19 (0.0152±0.0115)	28 (0.0098±0.0012)	2 (0.1428±0.0525)	12 (0.0236±0.0024)	5 (0.0499±0.0276)	7 (0.0412±0.0035)	22 (0.0131±0.0032)	14 (0.0175±0.0111)
ITALIAN	0.7934	21 (0.0085±0.0064)	22 (0.0085±0.0013)	2 (0.1317±0.0250)	10 (0.0213±0.0010)	13 (0.0150±0.0085)	4 (0.0587±0.0022)	26 (0.0076±0.0058)	34 (0.0070±0.0018)
JAPANESE	0.7578	24 (0.0090±0.0066)	36 (0.0054±0.0013)	2 (0.1288±0.0229)	10 (0.0257±0.0019)	15 (0.0147±0.0074)	4 (0.0591±0.0033)	35 (0.0054±0.0028)	16 (0.0144±0.0159)
MEXICAN	0.7881	18 (0.0119±0.0082)	32 (0.0088±0.0010)	2 (0.1266±0.0231)	12 (0.0211±0.0015)	10 (0.0240±0.0120)	4 (0.0540±0.0025)	25 (0.0097±0.0045)	21 (0.0107±0.0051)
SALAD	0.7782	28 (0.0075±0.0035)	30 (0.0072±0.0014)	2 (0.1267±0.0248)	11 (0.0218±0.0022)	10 (0.0250±0.0176)	4 (0.0582±0.0026)	22 (0.0084±0.0077)	36 (0.0058±0.0017)
AMERICAN-TRADITIONAL	0.7743	20 (0.0122±0.0090)	25 (0.0098±0.0040)	2 (0.1301±0.0319)	10 (0.0228±0.0017)	11 (0.0205±0.0140)	4 (0.0524±0.0021)	34 (0.0073±0.0012)	23 (0.0101±0.0040)
CHINESE	0.7413	23 (0.0104±0.0043)	34 (0.0080±0.0012)	2 (0.1402±0.0330)	11 (0.0201±0.0016)	10 (0.0227±0.0193)	5 (0.0522±0.0040)	29 (0.0088±0.0028)	21 (0.0110±0.0065)
COFFEE_TEA	0.8103	20 (0.0117±0.0129)	25 (0.0077±0.0017)	2 (0.1190±0.0334)	12 (0.0182±0.0018)	9 (0.0370±0.0216)	4 (0.0559±0.0026)	29 (0.0072±0.0069)	35 (0.0061±0.0013)
C0-_AMERICANTRADITIONAL_	0.8025	12 (0.0137±0.0118)	24 (0.0093±0.0015)	2 (0.1635±0.0504)	7 (0.0278±0.0033)	6 (0.0293±0.0234)	41 (0.0000±0.0000)	15 (0.0110±0.0068)	16 (0.0109±0.0025)
C0-_CHINESE_	0.818	14 (0.0120±0.0060)	18 (0.0115±0.0056)	2 (0.1864±0.0533)	7 (0.0203±0.0033)	6 (0.0210±0.0105)	41 (0.0000±0.0000)	16 (0.0117±0.0050)	15 (0.0117±0.0069)
C0-_COFFEE_TEA_	0.8109	14 (0.0107±0.0157)	20 (0.0091±0.0028)	2 (0.1566±0.0470)	7 (0.0205±0.0023)	6 (0.0328±0.0313)	41 (0.0000±0.0000)	26 (0.0076±0.0103)	27 (0.0073±0.0032)
C0-_FASTFOOD_	0.831	18 (0.0139±0.0124)	21 (0.0113±0.0023)	2 (0.1538±0.0683)	7 (0.0282±0.0045)	6 (0.0394±0.0333)	41 (0.0000±0.0000)	11 (0.0157±0.0045)	12 (0.0150±0.0037)
C0-_ITALIAN_	0.7847	12 (0.0115±0.0072)	13 (0.0112±0.0023)	2 (0.1616±0.0381)	6 (0.0251±0.0032)	7 (0.0209±0.0153)	41 (0.0000±0.0000)	20 (0.0091±0.0059)	25 (0.0075±0.0015)
C0-_JAPANESE_	0.7791	13 (0.0126±0.0078)	28 (0.0070±0.0020)	2 (0.1671±0.0372)	6 (0.0275±0.0046)	7 (0.0256±0.0218)	41 (0.0000±0.0000)	29 (0.0055±0.0032)	17 (0.0098±0.0087)
C0-_MEXICAN_	0.8131	15 (0.0119±0.0067)	23 (0.0102±0.0014)	2 (0.1719±0.0449)	6 (0.0243±0.0020)	7 (0.0231±0.0095)	35 (0.0000±0.0000)	12 (0.0125±0.0105)	24 (0.0101±0.0013)
C0-_SALAD_	0.764	24 (0.0092±0.0023)	15 (0.0100±0.0019)	2 (0.1565±0.0353)	6 (0.0256±0.0020)	7 (0.0234±0.0164)	35 (0.0000±0.0000)	22 (0.0095±0.0070)	27 (0.0082±0.0028)
C2-_AMERICANTRADITIONAL_	0.7073	28 (0.0132±0.0031)	24 (0.0135±0.0030)	2 (0.0766±0.0114)	12 (0.0234±0.0020)	13 (0.0214±0.0074)	4 (0.0721±0.0047)	30 (0.0129±0.0047)	19 (0.0140±0.0044)
C2-_CHINESE_	0.68	19 (0.0150±0.0057)	20 (0.0150±0.0033)	2 (0.0813±0.0092)	13 (0.0235±0.0034)	11 (0.0245±0.0104)	3 (0.0708±0.0075)	21 (0.0149±0.0050)	22 (0.0148±0.0037)
C2-_COFFEE_TEA_	0.7681	27 (0.0108±0.0065)	19 (0.0129±0.0035)	3 (0.0845±0.0136)	15 (0.0183±0.0028)	13 (0.0211±0.0089)	4 (0.0731±0.0061)	35 (0.0068±0.0039)	18 (0.0145±0.0085)
C2-_FASTFOOD_	0.7247	18 (0.0195±0.0049)	25 (0.0161±0.0026)	1 (0.0968±0.0223)	11 (0.0300±0.0039)	9 (0.0386±0.0138)	3 (0.0555±0.0050)	20 (0.0170±0.0025)	15 (0.0220±0.0072)
C2-_ITALIAN_	0.7692	17 (0.0138±0.0058)	18 (0.0138±0.0020)	4 (0.0782±0.0112)	11 (0.0237±0.0032)	14 (0.0207±0.0074)	3 (0.0799±0.0050)	32 (0.0088±0.0021)	31 (0.0093±0.0022)
C2-_JAPANESE_	0.6968	23 (0.0125±0.0064)	34 (0.0089±0.0026)	2 (0.0804±0.0122)	11 (0.0288±0.0047)	16 (0.0165±0.0034)	3 (0.0772±0.0068)	31 (0.0099±0.0049)	14 (0.0170±0.0147)
C2-_MEXICAN_	0.7284	16 (0.0163±0.0059)	27 (0.0133±0.0021)	3 (0.0716±0.0094)	13 (0.0230±0.0022)	14 (0.0212±0.0037)	4 (0.0713±0.0047)	30 (0.0120±0.0019)	19 (0.0159±0.0052)
C2-_SALAD_	0.7385	25 (0.0115±0.0036)	28 (0.0114±0.0022)	3 (0.0810±0.0103)	11 (0.0249±0.0030)	15 (0.0173±0.0076)	4 (0.0795±0.0084)	35 (0.0072±0.0028)	31 (0.0091±0.0019)
C3-_AMERICANTRADITIONAL_	0.7641	20 (0.0136±0.0052)	26 (0.0127±0.0016)	2 (0.1317±0.0373)	10 (0.0235±0.0023)	11 (0.0206±0.0060)	3 (0.0622±0.0049)	32 (0.0112±0.0040)	27 (0.0124±0.0021)
C3-_CHINESE_	0.761	20 (0.0137±0.0037)	22 (0.0135±0.0046)	1 (0.1516±0.0340)	10 (0.0202±0.0036)	11 (0.0196±0.0080)	4 (0.0638±0.0075)	31 (0.0111±0.0033)	24 (0.0128±0.0039)
C3-_COFFEE_TEA_	0.8042	19 (0.0128±0.0087)	22 (0.0120±0.0038)	2 (0.1194±0.0307)	14 (0.0192±0.0032)	10 (0.0241±0.0145)	3 (0.0745±0.0079)	33 (0.0083±0.0084)	25 (0.0110±0.0069)
C3-_FASTFOOD_	0.8045	14 (0.0198±0.0122)	28 (0.0112±0.0029)	1 (0.1575±0.0659)	12 (0.0257±0.0034)	7 (0.0371±0.0154)	5 (0.0428±0.0056)	19 (0.0155±0.0047)	17 (0.0172±0.0051)
C3-_ITALIAN_	0.7937	23 (0.0123±0.0049)	17 (0.0133±0.0019)	2 (0.1452±0.0292)	11 (0.0198±0.0020)	13 (0.0173±0.0049)	3 (0.0711±0.0066)	30 (0.0097±0.0015)	29 (0.0114±0.0032)
C3-_JAPANESE_	0.7972	21 (0.0118±0.0054)	31 (0.0088±0.0032)	2 (0.1193±0.0251)	10 (0.0216±0.0039)	12 (0.0195±0.0103)	3 (0.0765±0.0071)	32 (0.0085±0.0056)	34 (0.0081±0.0042)
C3-_MEXICAN_	0.8025	15 (0.0156±0.0076)	20 (0.0132±0.0023)	2 (0.1230±0.0266)	10 (0.0213±0.0020)	11 (0.0206±0.0077)	3 (0.0656±0.0044)	17 (0.0147±0.0104)	29 (0.0122±0.0033)
C3-_SALAD_	0.7993	24 (0.0107±0.0029)	23 (0.0112±0.0018)	2 (0.1305±0.0290)	11 (0.0237±0.0021)	10 (0.0275±0.0155)	3 (0.0740±0.0051)	30 (0.0090±0.0040)	28 (0.0098±0.0059)
C42-_MEXICAN_	0.8033	16 (0.0164±0.0084)	18 (0.0153±0.0028)	1 (0.1595±0.0446)	14 (0.0170±0.0038)	10 (0.0293±0.0194)	4 (0.0496±0.0082)	3 (0.0591±0.0253)	30 (0.0124±0.0030)
C9-_AMERICANTRADITIONAL_	0.7987	12 (0.0204±0.0159)	27 (0.0133±0.0030)	1 (0.1382±0.0518)	14 (0.0182±0.0026)	11 (0.0246±0.0125)	4 (0.0508±0.0070)	3 (0.0699±0.0411)	16 (0.0175±0.0159)
C9-_ITALIAN_	0.8178	19 (0.0161±0.0074)	13 (0.0175±0.0048)	1 (0.1523±0.0441)	12 (0.0203±0.0037)	10 (0.0240±0.0158)	4 (0.0524±0.0076)	3 (0.0561±0.0331)	27 (0.0137±0.0053)
C9-_MEXICAN_	0.8033	16 (0.0164±0.0084)	18 (0.0153±0.0028)	1 (0.1595±0.0446)	14 (0.0170±0.0038)	10 (0.0293±0.0194)	4 (0.0496±0.0082)	3 (0.0591±0.0253)	30 (0.0124±0.0030)

Subset	Accuracy	c42fav	social1rating	sanitation	c27fav	c54fav	amount	place	kolreview
Full Data	0.7917	26 (0.0119±0.0042)	15 (0.0155±0.0007)	34 (0.0078±0.0005)	19 (0.0127±0.0027)	27 (0.0119±0.0018)	41 (0.0006±0.0001)	33 (0.0087±0.0022)	35 (0.0059±0.0008)
community0	0.814	21 (0.0123±0.0035)	36 (0.0000±0.0000)	27 (0.0076±0.0009)	14 (0.0133±0.0011)	22 (0.0122±0.0014)	34 (0.0005±0.0001)	23 (0.0122±0.0047)	37 (0.0000±0.0000)
community13	0.8135	21 (0.0178±0.0035)	25 (0.0146±0.0025)	33 (0.0077±0.0011)	17 (0.0186±0.0021)	13 (0.0196±0.0019)	39 (0.0009±0.0004)	31 (0.0107±0.0020)	35 (0.0063±0.0015)
community19	0.7866	25 (0.0159±0.0033)	30 (0.0126±0.0023)	35 (0.0053±0.0014)	20 (0.0189±0.0030)	16 (0.0199±0.0044)	39 (0.0009±0.0004)	32 (0.0096±0.0022)	34 (0.0054±0.0018)
community2	0.7285	27 (0.0159±0.0019)	8 (0.0398±0.0018)	33 (0.0104±0.0005)	20 (0.0168±0.0016)	18 (0.0170±0.0023)	40 (0.0010±0.0001)	34 (0.0090±0.0013)	22 (0.0165±0.0014)
community3	0.8025	27 (0.0145±0.0029)	26 (0.0150±0.0010)	33 (0.0076±0.0007)	16 (0.0164±0.0010)	20 (0.0157±0.0012)	40 (0.0005±0.0001)	32 (0.0112±0.0029)	35 (0.0058±0.0007)
community42	0.8059	1 (0.1466±0.0539)	32 (0.0093±0.0015)	33 (0.0076±0.0012)	16 (0.0180±0.0032)	15 (0.0183±0.0030)	40 (0.0002±0.0002)	28 (0.0131±0.0038)	36 (0.0028±0.0006)
community5	0.8086	25 (0.0159±0.0031)	28 (0.0144±0.0024)	35 (0.0065±0.0015)	22 (0.0189±0.0024)	17 (0.0197±0.0022)	40 (0.0013±0.0006)	32 (0.0110±0.0033)	34 (0.0075±0.0011)
community9	0.8356	23 (0.0163±0.0054)	27 (0.0128±0.0014)	34 (0.0067±0.0009)	14 (0.0187±0.0039)	18 (0.0169±0.0020)	40 (0.0002±0.0001)	32 (0.0094±0.0024)	35 (0.0038±0.0006)
FASTFOOD	0.7781	15 (0.0174±0.0093)	27 (0.0102±0.0010)	21 (0.0142±0.0022)	24 (0.0114±0.0053)	29 (0.0092±0.0023)	40 (0.0001±0.0001)	34 (0.0066±0.0017)	35 (0.0048±0.0008)
ITALIAN	0.7934	35 (0.0068±0.0030)	11 (0.0191±0.0011)	36 (0.0066±0.0004)	20 (0.0092±0.0034)	33 (0.0071±0.0012)	41 (0.0001±0.0000)	18 (0.0099±0.0018)	32 (0.0073±0.0011)
JAPANESE	0.7578	37 (0.0048±0.0010)	12 (0.0191±0.0017)	23 (0.0096±0.0011)	26 (0.0084±0.0049)	30 (0.0072±0.0014)	41 (0.0002±0.0001)	19 (0.0117±0.0020)	25 (0.0086±0.0008)
MEXICAN	0.7881	22 (0.0102±0.0058)	13 (0.0159±0.0009)	28 (0.0097±0.0007)	27 (0.0097±0.0011)	20 (0.0107±0.0044)	41 (0.0002±0.0000)	29 (0.0090±0.0021)	36 (0.0065±0.0006)
SALAD	0.7782	34 (0.0063±0.0029)	12 (0.0188±0.0015)	19 (0.0096±0.0009)	20 (0.0087±0.0035)	25 (0.0079±0.0017)	40 (0.0009±0.0002)	17 (0.0111±0.0020)	27 (0.0076±0.0010)
AMERICAN-TRADITIONAL	0.7743	26 (0.0097±0.0049)	13 (0.0167±0.0007)	31 (0.0086±0.0007)	27 (0.0095±0.0020)	16 (0.0134±0.0123)	41 (0.0005±0.0001)	22 (0.0114±0.0030)	36 (0.0067±0.0008)
CHINESE	0.7413	31 (0.0087±0.0026)	13 (0.0156±0.0014)	16 (0.0123±0.0011)	32 (0.0085±0.0024)	26 (0.0091±0.0026)	41 (0.0001±0.0001)	18 (0.0116±0.0036)	36 (0.0065±0.0009)
COFFEE_TEA	0.8103	23 (0.0080±0.0049)	11 (0.0195±0.0020)	26 (0.0075±0.0009)	34 (0.0062±0.0032)	17 (0.0122±0.0115)	40 (0.0016±0.0003)	22 (0.0095±0.0021)	27 (0.0074±0.0014)
C0-_AMERICANTRADITIONAL_	0.8025	25 (0.0092±0.0015)	36 (0.0000±0.0000)	28 (0.0074±0.0011)	18 (0.0107±0.0018)	11 (0.0145±0.0137)	34 (0.0005±0.0002)	9 (0.0168±0.0043)	35 (0.0000±0.0000)
C0-_CHINESE_	0.818	26 (0.0098±0.0025)	35 (0.0000±0.0000)	10 (0.0150±0.0034)	25 (0.0100±0.0018)	20 (0.0109±0.0032)	40 (0.0000±0.0000)	9 (0.0154±0.0045)	34 (0.0000±0.0000)
C0-_COFFEE_TEA_	0.8109	25 (0.0080±0.0039)	36 (0.0000±0.0000)	22 (0.0089±0.0012)	21 (0.0091±0.0072)	11 (0.0125±0.0117)	31 (0.0023±0.0006)	9 (0.0134±0.0029)	35 (0.0000±0.0000)
C0-_FASTFOOD_	0.831	9 (0.0200±0.0085)	38 (0.0000±0.0000)	19 (0.0133±0.0032)	15 (0.0142±0.0097)	23 (0.0096±0.0022)	33 (0.0001±0.0002)	27 (0.0071±0.0019)	37 (0.0000±0.0000)
C0-_ITALIAN_	0.7847	26 (0.0067±0.0016)	39 (0.0000±0.0000)	28 (0.0060±0.0012)	14 (0.0110±0.0036)	24 (0.0075±0.0018)	34 (0.0000±0.0000)	10 (0.0146±0.0027)	38 (0.0000±0.0000)
C0-_JAPANESE_	0.7791	22 (0.0089±0.0087)	35 (0.0000±0.0000)	16 (0.0108±0.0028)	15 (0.0109±0.0076)	18 (0.0090±0.0033)	40 (0.0000±0.0000)	11 (0.0141±0.0039)	36 (0.0000±0.0000)
C0-_MEXICAN_	0.8131	21 (0.0105±0.0025)	37 (0.0000±0.0000)	26 (0.0098±0.0014)	19 (0.0110±0.0015)	13 (0.0121±0.0036)	34 (0.0001±0.0001)	11 (0.0131±0.0034)	38 (0.0000±0.0000)
C0-_SALAD_	0.764	20 (0.0097±0.0048)	38 (0.0000±0.0000)	16 (0.0100±0.0015)	18 (0.0099±0.0022)	17 (0.0100±0.0024)	33 (0.0007±0.0002)	10 (0.0146±0.0024)	37 (0.0000±0.0000)
C2-_AMERICANTRADITIONAL_	0.7073	31 (0.0123±0.0022)	7 (0.0429±0.0028)	33 (0.0117±0.0013)	21 (0.0139±0.0018)	17 (0.0176±0.0073)	40 (0.0002±0.0002)	26 (0.0134±0.0028)	16 (0.0189±0.0016)
C2-_CHINESE_	0.68	18 (0.0158±0.0031)	8 (0.0411±0.0047)	16 (0.0171±0.0024)	30 (0.0136±0.0034)	26 (0.0140±0.0044)	39 (0.0000±0.0000)	34 (0.0108±0.0024)	15 (0.0191±0.0040)
C2-_COFFEE_TEA_	0.7681	25 (0.0110±0.0045)	7 (0.0492±0.0048)	32 (0.0089±0.0024)	28 (0.0105±0.0040)	16 (0.0156±0.0093)	39 (0.0019±0.0009)	22 (0.0120±0.0028)	14 (0.0196±0.0038)
C2-_FASTFOOD_	0.7247	16 (0.0199±0.0071)	13 (0.0272±0.0035)	22 (0.0167±0.0041)	27 (0.0148±0.0028)	29 (0.0139±0.0026)	39 (0.0003±0.0003)	34 (0.0060±0.0022)	26 (0.0149±0.0021)
C2-_ITALIAN_	0.7692	30 (0.0099±0.0054)	7 (0.0478±0.0039)	33 (0.0087±0.0016)	19 (0.0137±0.0030)	27 (0.0116±0.0024)	40 (0.0004±0.0000)	29 (0.0108±0.0024)	15 (0.0191±0.0026)
C2-_JAPANESE_	0.6968	35 (0.0086±0.0027)	8 (0.0472±0.0060)	27 (0.0110±0.0026)	28 (0.0109±0.0039)	18 (0.0139±0.0036)	39 (0.0006±0.0004)	20 (0.0129±0.0022)	12 (0.0211±0.0045)
C2-_MEXICAN_	0.7284	24 (0.0140±0.0055)	8 (0.0414±0.0027)	32 (0.0109±0.0017)	20 (0.0147±0.0024)	17 (0.0163±0.0016)	40 (0.0001±0.0001)	34 (0.0102±0.0017)	15 (0.0183±0.0026)
C2-_SALAD_	0.7385	22 (0.0120±0.0061)	7 (0.0480±0.0038)	21 (0.0120±0.0021)	19 (0.0137±0.0037)	20 (0.0122±0.0019)	39 (0.0015±0.0005)	26 (0.0114±0.0022)	13 (0.0207±0.0038)
C3-_AMERICANTRADITIONAL_	0.7641	31 (0.0116±0.0067)	14 (0.0169±0.0021)	33 (0.0077±0.0011)	25 (0.0127±0.0017)	17 (0.0144±0.0088)	40 (0.0004±0.0002)	16 (0.0152±0.0041)	34 (0.0066±0.0012)
C3-_CHINESE_	0.761	23 (0.0132±0.0043)	21 (0.0136±0.0034)	27 (0.0121±0.0036)	29 (0.0117±0.0031)	25 (0.0127±0.0030)	41 (0.0000±0.0000)	14 (0.0157±0.0057)	35 (0.0066±0.0019)
C3-_COFFEE_TEA_	0.8042	30 (0.0088±0.0070)	13 (0.0192±0.0040)	29 (0.0089±0.0020)	34 (0.0078±0.0027)	24 (0.0114±0.0066)	39 (0.0017±0.0009)	21 (0.0123±0.0037)	35 (0.0076±0.0018)
C3-_FASTFOOD_	0.8045	15 (0.0194±0.0064)	32 (0.0084±0.0026)	26 (0.0115±0.0025)	20 (0.0141±0.0055)	23 (0.0126±0.0048)	41 (0.0000±0.0000)	31 (0.0086±0.0029)	36 (0.0034±0.0014)
C3-_ITALIAN_	0.7937	32 (0.0091±0.0027)	12 (0.0179±0.0028)	36 (0.0056±0.0011)	15 (0.0145±0.0047)	25 (0.0119±0.0026)	40 (0.0000±0.0000)	24 (0.0120±0.0022)	34 (0.0077±0.0016)
C3-_JAPANESE_	0.7972	29 (0.0090±0.0036)	13 (0.0184±0.0037)	27 (0.0096±0.0016)	23 (0.0104±0.0041)	18 (0.0128±0.0040)	40 (0.0000±0.0000)	15 (0.0146±0.0034)	30 (0.0088±0.0027)
C3-_MEXICAN_	0.8025	19 (0.0138±0.0070)	13 (0.0169±0.0015)	33 (0.0082±0.0015)	21 (0.0131±0.0020)	16 (0.0150±0.0035)	40 (0.0004±0.0002)	32 (0.0107±0.0030)	35 (0.0061±0.0010)
C3-_SALAD_	0.7993	32 (0.0083±0.0018)	14 (0.0178±0.0021)	31 (0.0086±0.0018)	22 (0.0112±0.0025)	20 (0.0122±0.0020)	40 (0.0004±0.0002)	17 (0.0141±0.0028)	34 (0.0078±0.0014)
C42-_MEXICAN_	0.8033	27 (0.0131±0.0034)	29 (0.0127±0.0024)	31 (0.0123±0.0023)	13 (0.0180±0.0047)	21 (0.0144±0.0046)	41 (0.0000±0.0000)	33 (0.0096±0.0029)	36 (0.0030±0.0014)
C9-_AMERICANTRADITIONAL_	0.7987	17 (0.0171±0.0135)	29 (0.0123±0.0032)	35 (0.0043±0.0020)	21 (0.0159±0.0042)	19 (0.0171±0.0092)	40 (0.0000±0.0002)	32 (0.0108±0.0031)	34 (0.0050±0.0017)
C9-_ITALIAN_	0.8178	29 (0.0106±0.0039)	28 (0.0125±0.0036)	36 (0.0043±0.0020)	21 (0.0158±0.0080)	23 (0.0146±0.0054)	40 (0.0000±0.0000)	31 (0.0102±0.0034)	34 (0.0055±0.0020)
C9-_MEXICAN_	0.8033	27 (0.0131±0.0034)	29 (0.0127±0.0024)	31 (0.0123±0.0023)	13 (0.0180±0.0047)	21 (0.0144±0.0046)	41 (0.0000±0.0000)	33 (0.0096±0.0029)	36 (0.0030±0.0014)

Subset	Accuracy	c2fav
Full Data	0.7917	20 (0.0124±0.0053)
community0	0.814	26 (0.0094±0.0035)
community13	0.8135	26 (0.0141±0.0037)
community19	0.7866	28 (0.0129±0.0048)
community2	0.7285	7 (0.0405±0.0151)
community3	0.8025	15 (0.0164±0.0115)
community42	0.8059	29 (0.0130±0.0046)
community5	0.8086	23 (0.0189±0.0143)
community9	0.8356	29 (0.0125±0.0051)
FASTFOOD	0.7781	16 (0.0171±0.0097)
ITALIAN	0.7934	29 (0.0074±0.0035)
JAPANESE	0.7578	17 (0.0127±0.0116)
MEXICAN	0.7881	17 (0.0133±0.0073)
SALAD	0.7782	18 (0.0096±0.0053)
AMERICAN-TRADITIONAL	0.7743	19 (0.0122±0.0092)
CHINESE	0.7413	24 (0.0099±0.0051)
COFFEE_TEA	0.8103	13 (0.0162±0.0131)
C0-_AMERICANTRADITIONAL_	0.8025	27 (0.0075±0.0077)
C0-_CHINESE_	0.818	29 (0.0076±0.0044)
C0-_COFFEE_TEA_	0.8109	17 (0.0097±0.0125)
C0-_FASTFOOD_	0.831	13 (0.0147±0.0106)
C0-_ITALIAN_	0.7847	27 (0.0065±0.0037)
C0-_JAPANESE_	0.7791	23 (0.0085±0.0059)
C0-_MEXICAN_	0.8131	22 (0.0104±0.0052)
C0-_SALAD_	0.764	28 (0.0075±0.0041)
C2-_AMERICANTRADITIONAL_	0.7073	11 (0.0271±0.0116)
C2-_CHINESE_	0.68	12 (0.0243±0.0109)
C2-_COFFEE_TEA_	0.7681	9 (0.0329±0.0133)
C2-_FASTFOOD_	0.7247	7 (0.0494±0.0206)
C2-_ITALIAN_	0.7692	13 (0.0217±0.0110)
C2-_JAPANESE_	0.6968	13 (0.0200±0.0145)
C2-_MEXICAN_	0.7284	7 (0.0477±0.0213)
C2-_SALAD_	0.7385	12 (0.0235±0.0112)
C3-_AMERICANTRADITIONAL_	0.7641	18 (0.0144±0.0135)
C3-_CHINESE_	0.761	18 (0.0138±0.0110)
C3-_COFFEE_TEA_	0.8042	18 (0.0129±0.0141)
C3-_FASTFOOD_	0.8045	25 (0.0125±0.0076)
C3-_ITALIAN_	0.7937	31 (0.0092±0.0031)
C3-_JAPANESE_	0.7972	36 (0.0078±0.0076)
C3-_MEXICAN_	0.8025	22 (0.0129±0.0094)
C3-_SALAD_	0.7993	27 (0.0102±0.0074)
C42-_MEXICAN_	0.8033	19 (0.0149±0.0075)
C9-_AMERICANTRADITIONAL_	0.7987	30 (0.0121±0.0095)
C9-_ITALIAN_	0.8178	30 (0.0103±0.0054)
C9-_MEXICAN_	0.8033	19 (0.0149±0.0075)