Rest o Recommend

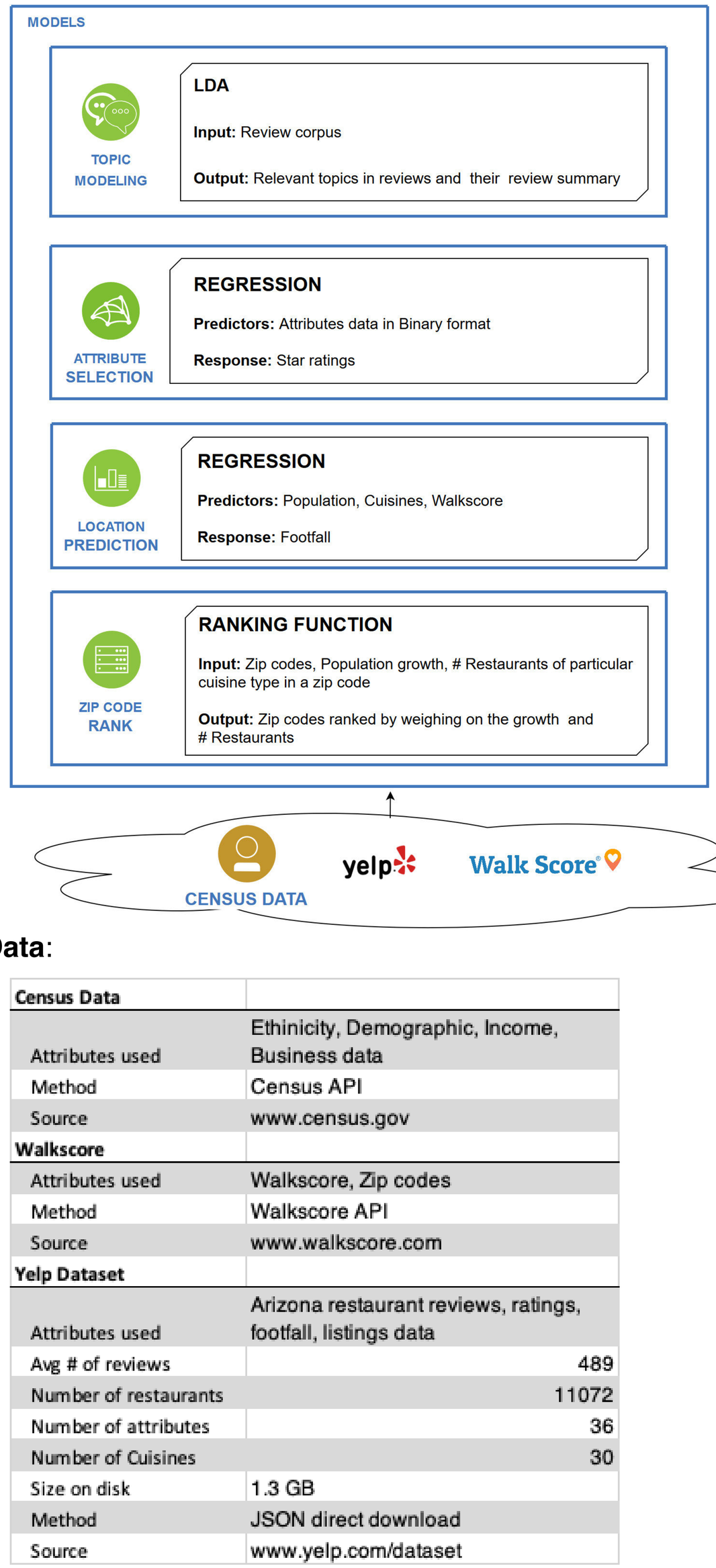
Nirupam I Kireeti I Navya I Etienne I Xiu I Prashanth

Motivation :

Many dream of owning a restaurant. However, restaurant business is hard. According to CNBC, 60% of restaurants fail in first year and 80% of restaurants fail before their fifth year. Main reasons of failure are: poor location choice, unable to get actionable insights from customer feedback and lack of competitive intelligence. We want to help restaurant owners use data to make decisions, thereby increasing their chance of success.

Approach :

Currently, restaurant owners rely on experience, intuition or approach consulting firms such as restaurantowner.com to run their business. Our approach, when compared to others, is as follows:



Data:

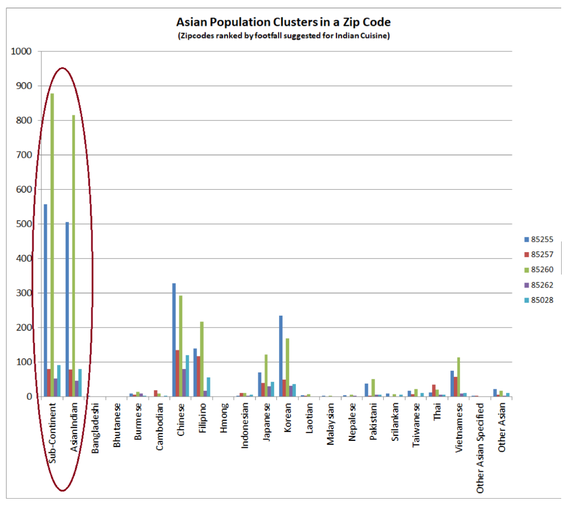
Census Data	
Attributes used	Ethinicity, Demographic, Income, Business data
Method	Census API
Source	www.census.gov
Walkscore	
Attributes used	Walkscore, Zip codes
Method	Walkscore API
Source	www.walkscore.com
Yelp Dataset	
Attributes used	Arizona restaurant reviews, ratings, footfall, listings data
Avg # of reviews	489
Number of restaurants	11072
Number of attributes	36
Number of Cuisines	30
Size on disk	1.3 GB
Method	JSON direct download
Source	www.yelp.com/dataset

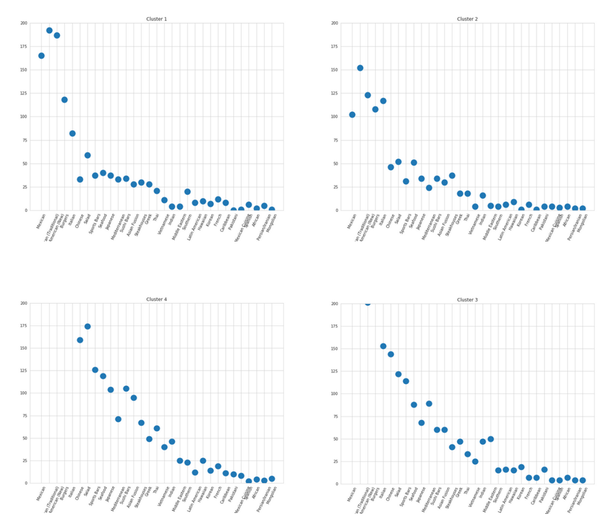
Experiments and Results :

Location Prediction:

We obtained an R-Square of 42.47%, RMSE of 2.67 for footfall using XGBoost Regressor. We observed that the zip codes predicted for higher footfall have a good representation of the ethnic groups that are associated with the selected cuisine

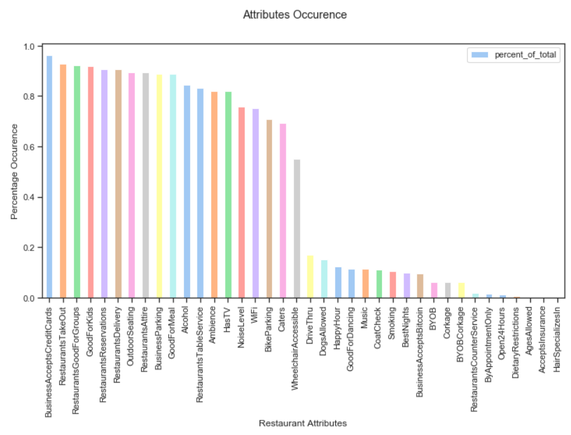
Model	RMSE	R-Square
Linear regression	1.7	24.19%
Random Forrest	2.72	38.20%
XGBoost	2.67	42.47%

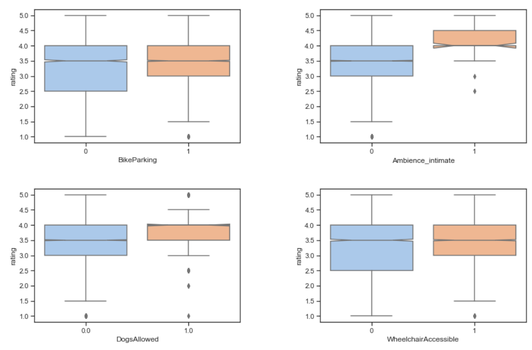




Attributes Evaluation:

We evaluated whether customers care about the attributes predicted by our model. We mined through reviews for 25 restaurants to evaluate the topics suggested by our model and observed the following prediction accuracy:



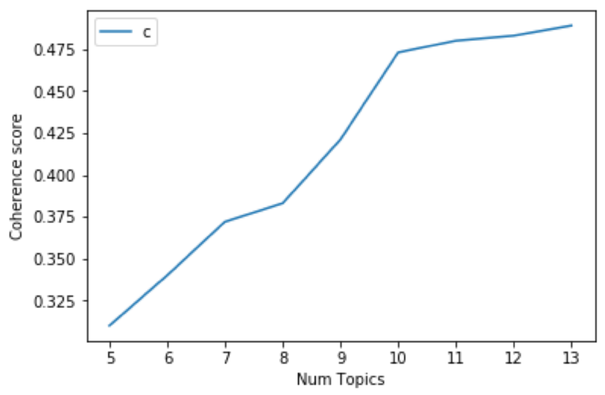


Percent of Cases	Result
56% of cases	3 or more suggestions were relevant
12% of cases	2 out of 5 suggestions were relevant
8% of cases	1 out of 5 suggestions were relevant
24% of cases	None of the suggestions were relevant

Attribute	Sample review excerpt
'Ambience'	This place feels a bit out-of-the way, with a small interior and not a whole lot to the ambiance. The sushi, however, is superb.

Text Analysis:

To determine the best number of topics to be retrieved from the reviews we computed topic coherence metric with different number of topics and obtained the optimal topic coherence as 10.



Summary Review: All and all the place is a great burger at a decent price. I would certainly recommend.

Conclusion:

The closest work related to our project is done by Lu et.al. But the work mainly concentrated on predicting whether an existing restaurant would be successful or not based on the yelp reviews. Our work is novel as we tried to suggest restaurant location where it would earn more revenues based on footfall metric. We later used LDA for topic modeling and TextRank algorithm for automatic summarization. Our model is unique as it combines various models and gives a one-stop solution for various restaurant setup and improvement requirements.