



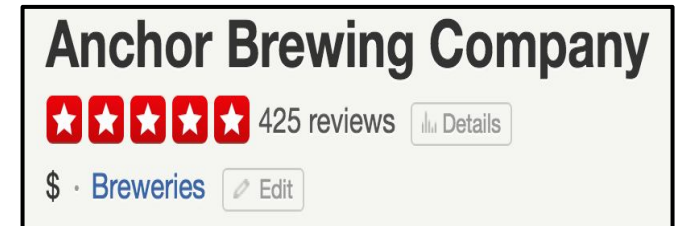
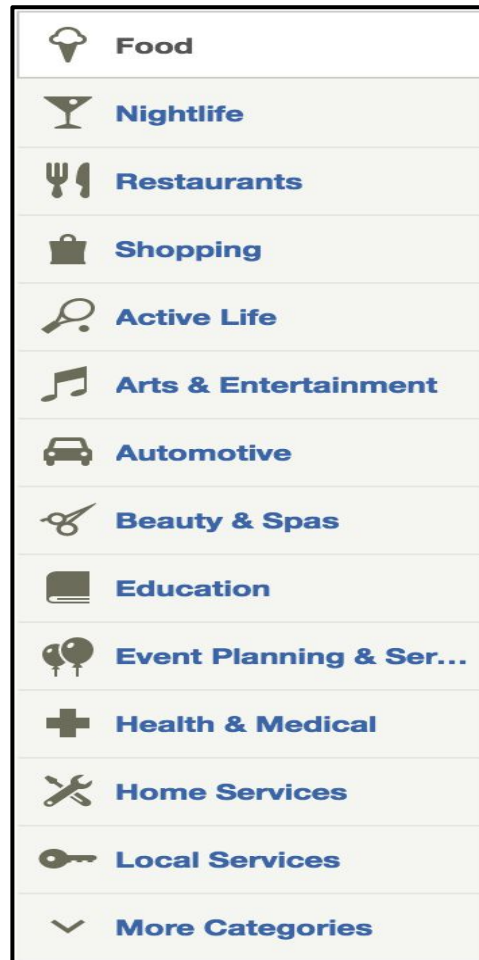
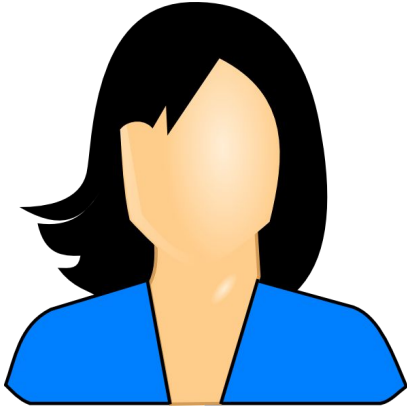
Scalable Recommendation System for Yelp

Piyush Bhargava, Sakshi Bhargava, Chhavi Choudhury



Project Overview

Yelp Business Categories



Data

kaggle RecSys2013: Yelp Business Rating Prediction

Recommendations for Businesses in Phoenix, AZ with  S3 &  EC2

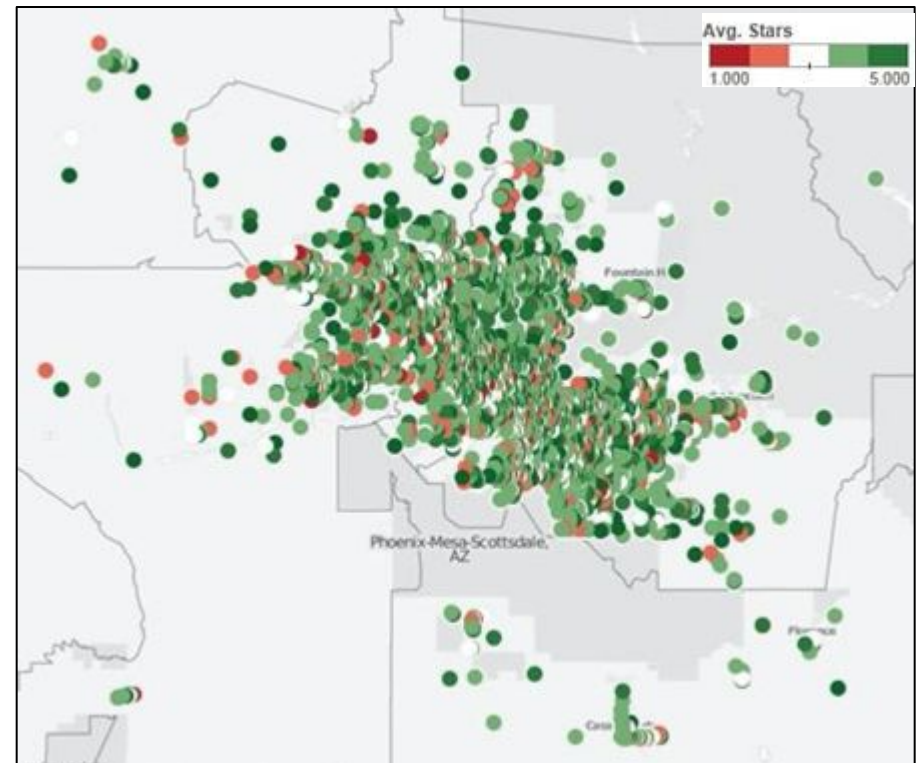
14.3K Businesses

53.3K Users






























266K Reviews

10K Check-Ins

	Known User	Unknown User
Known Business	33.2%	41.1%
Unknown Business	11.2%	14.5%

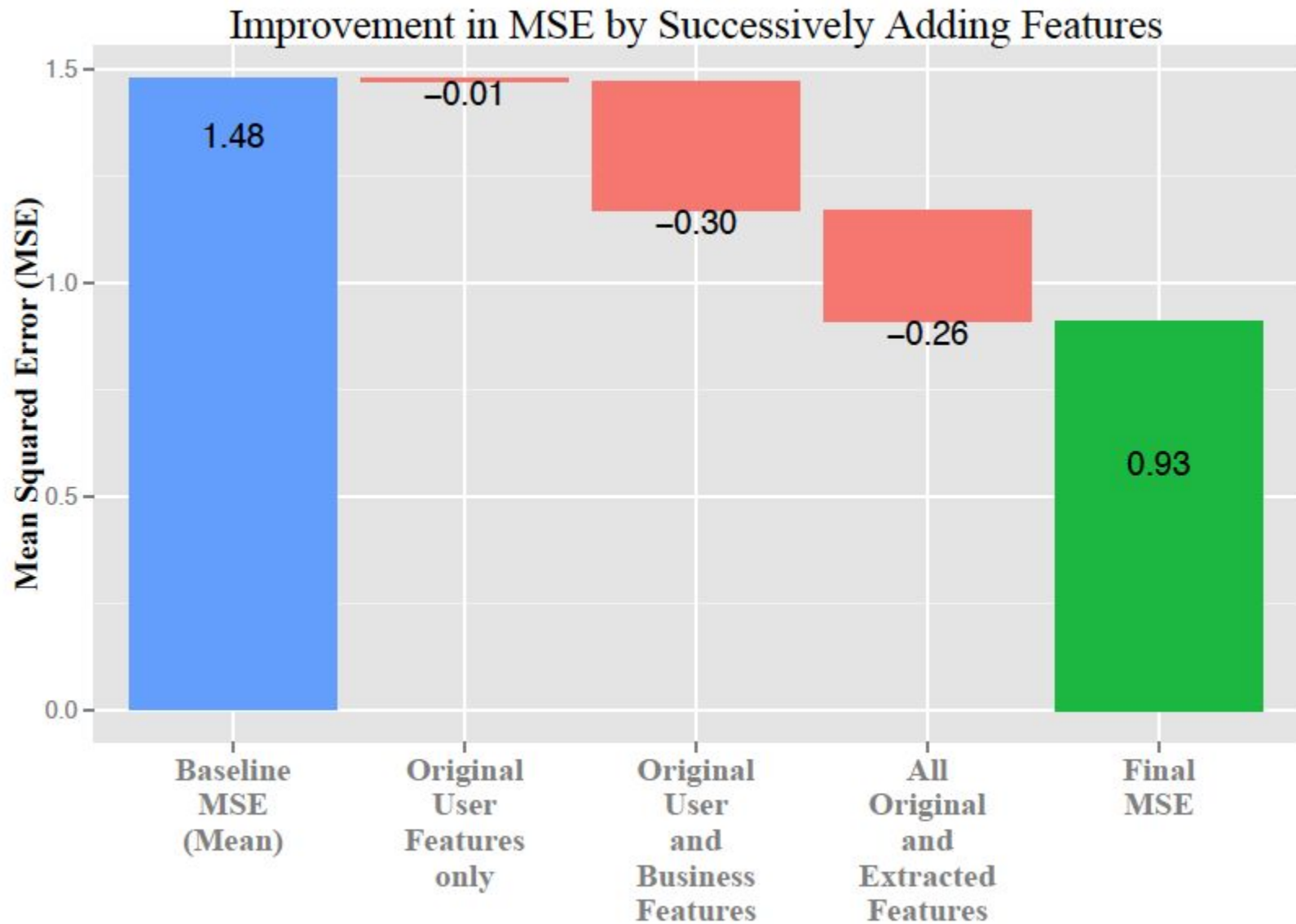


Feature Engineering by Segments

		Segment 1	Segment 2	Segment 3	Segment 4	Segment 5
Business Features	Average Star Ratings					
	Categories*					
	Review Count					
	Avg. Ratings for Clustered Locations*					
User features	Average Star Ratings					
	Review Count					
	User Gender*					
Check-In Features	Weekday /Weekend*					
Review	User Votes*					

* Extracted features

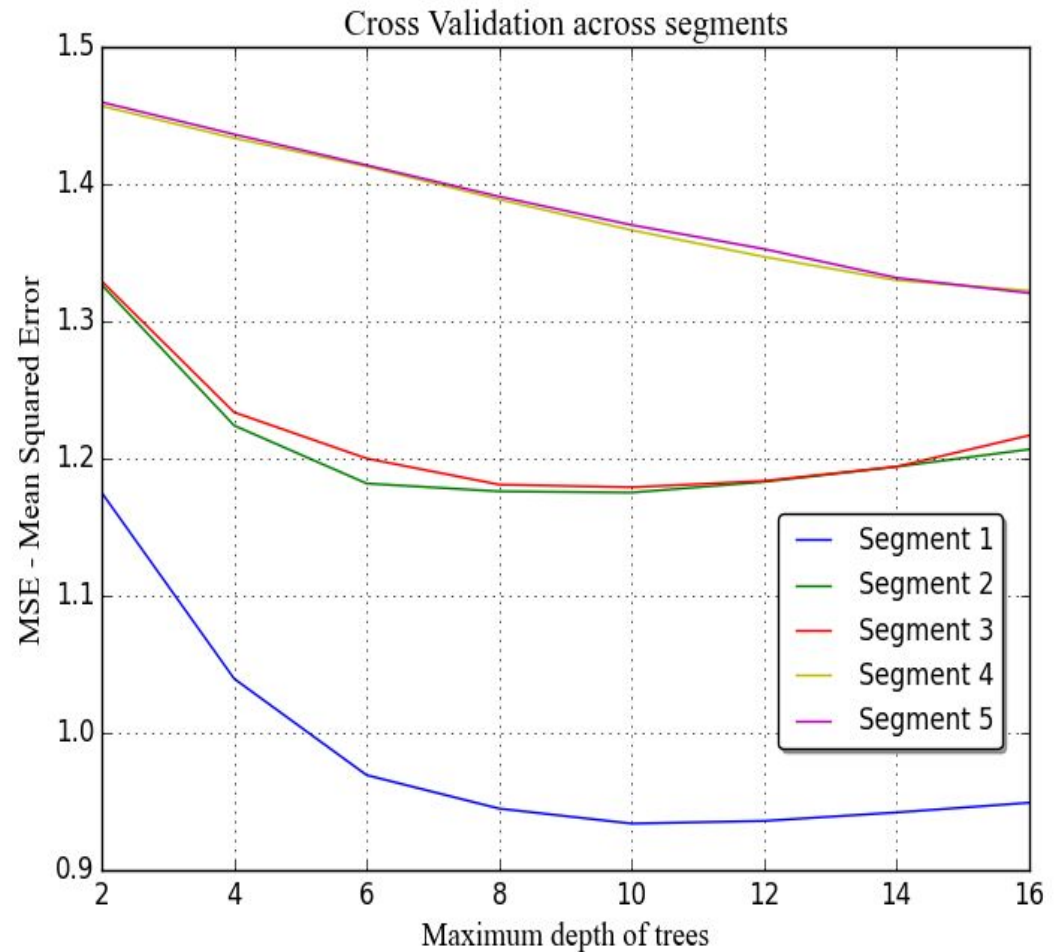
Feature Impact Waterfall Graph



Model Selection

Method	RMSE
Random Forest	1.267
Random Forest + Matrix Factorisation	1.275
Gradient Boosting	1.278
Matrix Factorisation	1.291
Logistic Regression	1.298
Linear Regression	1.290
Global Mean Baseline	1.390

Kaggle 1st Rank RMSE	1.212
----------------------	--------------



Random Forest - Cross Validation

Thank You



Chhavi Choudhury



chhavichoudhury@gmail.com



[chhavichoudhury](https://www.linkedin.com/in/chhavichoudhury)



Piyush Bhargava



piyush.bhargava.07@gmail.com



[piyush-bhargava](https://www.linkedin.com/in/piyush-bhargava)



Sakshi Bhargava



sakshibhargava024@gmail.com



[sbhargava024](https://www.linkedin.com/in/sbhargava024)