





Problem: Sentiment
Classification of London top
300 restaurant reviews and
building recommendation
system via restaurant
description, summary and
user comments.

Solution: To collect data from tripadvisor.com and save the data to database,MongoDB.

















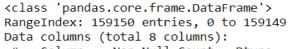




2815

restaurant





	Data	corumns (rotal & columns):				
	#	Column	Non-Null Count	Dtype			
	0	name	159150 non-null	object			
	1	rating	159150 non-null	float64			
	2	comments	159150 non-null	object			
	3	users	159150 non-null	object			
	4	summary	52641 non-null	object			
	5	info	159150 non-null	object			
	6	street	159150 non-null	object			
	7	postCode	159150 non-null	object			
dtypes: float64(1), object(7)							
	memory usage: 9.7+ MB						





113,303













DATA								
name	rating	comments	users	summary	info	street	postCode	-
London Steakhouse Co.	3.0	There was only 3 choices of starters and was n	natspinks2018	The London Steakhouse Company is a traditional	TRY 229 - TRY 1,177 Vegetarian Friendly, Glute	University House 109-117 Middlesex Street Devo	London E1 7JF England	A
London Steakhouse Co.	1.0	We were stared at as we walked in then had to	kellieblake2018	The London Steakhouse Company is a traditional	TRY 229 - TRY 1,177 Vegetarian Friendly, Glute	University House 109-117 Middlesex Street Devo	London E1 7JF England	WE
London Steakhouse Co.	5.0	Great food and good wine with great service th	Lidstone	The London Steakhouse Company is a traditional	TRY 229 - TRY 1,177 Vegetarian Friendly, Glute	University House 109-117 Middlesex Street Devo	London E1 7JF England	
London Steakhouse Co.	3.0	Every-ones taste is different, the food was ok	beenonhols	The London Steakhouse Company is a traditional	TRY 229 - TRY 1,177 Vegetarian Friendly, Glute	University House 109-117 Middlesex Street Devo	London E1 7JF England	
London Steakhouse Co.	1.0	My wife and I visited this restaurant as part	barls-shed21	The London Steakhouse Company is a traditional	TRY 229 - TRY 1,177 Vegetarian Friendly, Glute	University House 109-117 Middlesex Street Devo	London E1 7JF England	A







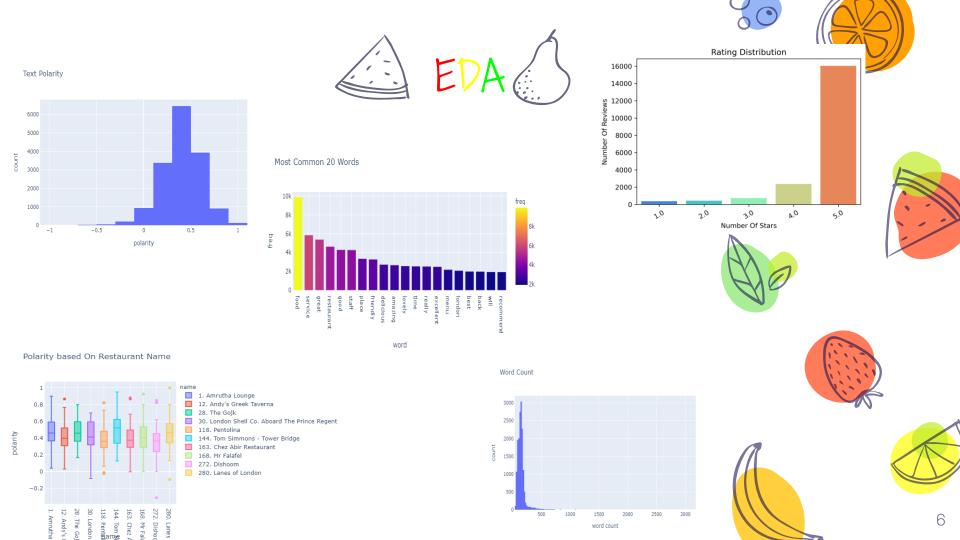


DATA & METHODOLOGY

- Data Collection and store DB
 - Preprocess the Text & EDA
 - Classification
 - Recommendation























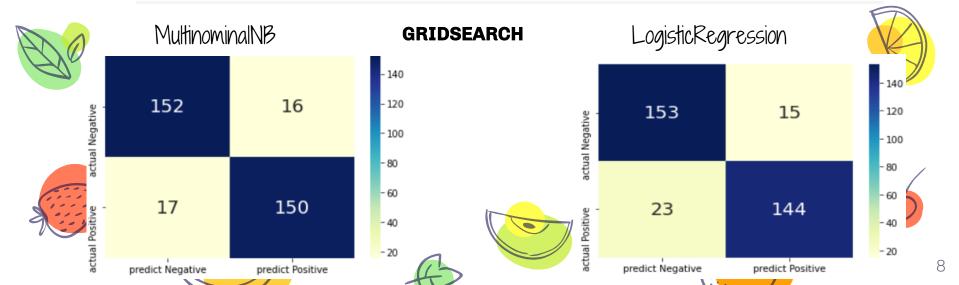








Model	Time x 1000	Train Accuracy	Test Accuracy	Precision	recall	f1	CV Accuracy	CV f1
LogisticRegression	16.031504	0.970475	0.895522	0.907407	0.880240	0.893617	0.867848	0.865848
GradientBoostingClassifier	1123.633862	0.961489	0.859701	0.905405	0.802395	0.850794	0.812686	0.802783
MultinomialNB	5.143166	0.975610	0.901493	0.903614	0.898204	0.900901	0.867848	0.868237
BernoulliNB	5.001068	0.953787	0.871642	0.840659	0.916168	0.876791	0.840891	0.855548
RandomForestClassifier	840.375185	1.000000	0.850746	0.882353	0.808383	0.843750	0.854993	0.840287
XGBClassifier	413.738489	0.998716	0.862687	0.885350	0.832335	0.858025	0.828003	0.820826
LinearSVC	4.000187	1.000000	0.865672	0.867470	0.862275	0.864865	0.857591	0.857599

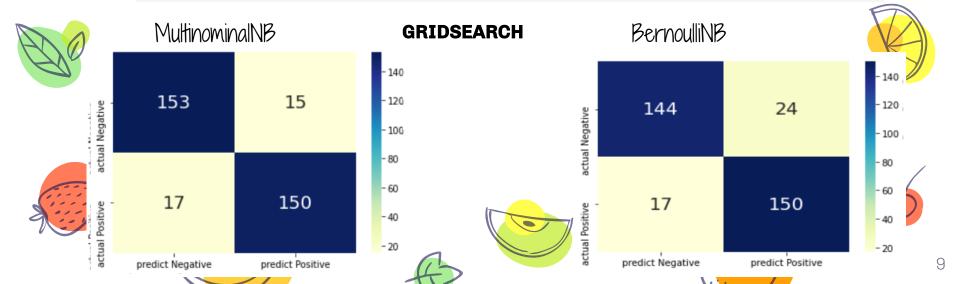






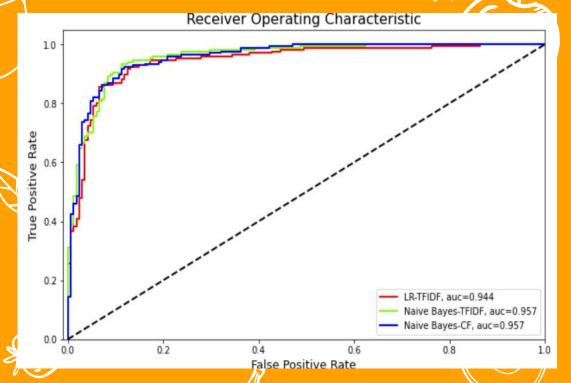


	Model	Time x 1000	Train Accuracy	Test Accuracy	Precision	recall	f1	CV Accuracy	CV f1
0	LogisticRegression	4.000425	1.000000	0.865672	0.876543	0.850299	0.863222	0.839642	0.840161
1	GradientBoostingClassifier	447.888374	0.931964	0.844776	0.870968	0.808383	0.838509	0.822976	0.809923
2	MultinomialNB	1.998901	0.971759	0.904478	0.909091	0.898204	0.903614	0.860088	0.859671
3	BernoulliNB	1.000643	0.953787	0.871642	0.840659	0.916168	0.876791	0.840891	0.855548
4	RandomForestClassifier	493.089676	1.000000	0.889552	0.911392	0.862275	0.886154	0.852463	0.849941
5	XGBClassifier	117.519855	0.988447	0.850746	0.854545	0.844311	0.849398	0.824224	0.821063
6	LinearSVC	9.000778	1.000000	0.835821	0.833333	0.838323	0.835821	0.837011	0.838157



© ROC Curve





Weight?	Feature
+0.586	love
+0.473	great
+0.395	excel
+0.392	amaz
+0.380	delici
+0.338	friend
+0.305	recommend
+0.298	thank
+0.290	fantast
+0.241	best
+0.238	servic
	ore positive
1227 mo	re negative
-0.241	bland
-0.260	terribl
-0.267	averag
-0.269	rude
-0.286	cold
-0.287	poor
-0.298	tabl
-0.409	ask
-0.645	order











Based on user comments

get n recommendations(" The Lounge Cafe", cosine sim ,5)

```
' Cafe Phillies',
 Bisou',
 Shepherdess Cafe',
' Cafe Continente',
```

Cafe Parisienne']

Based on summary, info and user comments



- ' Cafe Maya',
 - Liz Cafe',
 - ' Dulce Cafe',
 - ' Cafe Parisienne',
 - ' Cafe Parisienne']





































