

# UI-case1

August 2, 2020

## 1 UI - Case 1 for Yelp User Modeling

## 2 - What Drives User to Give high-star ( $\geq 4$ stars) reviews?

Before started this visualization, you shall have generated data by following this ReadMe file  
ReadMe file: <https://github.com/hycinthgeo/YelpUserModeling>

```
In [1]: from joblib import load
import sys
sys.path.append('src/python')
import modeling as train
from main import *

# parameter that UI user are encouraged to control
coef_topK = 50 #top coefficients with large magnitude

# please not change the following
App = myapp()
io_config = App.init_IO()
logger = App.init_logging(io_config["log path"], "UI-case1.log")
case1_model_artefact_path = io_config["modelArtefact path"]
feature_names_path = io_config["featureNames path"]

# variables to display graph
feature_names = list(pd.read_csv(feature_names_path, sep=" ", header=None)[0])
clf = load(case1_model_artefact_path)
coef = clf.coef_[0]
fig_path_coef = ""

# variables to prepare topK-coefficient table
coef_df = pd.DataFrame(coef, columns=['coef'])
coef_df['feature'] = feature_names
coef_df['abs(coef)'] = coef_df['coef'].apply(lambda x: abs(x))
coef_topK_df = coef_df.sort_values(['abs(coef)'], ascending=False).reset_index().loc[:coef_topK]
```

## 2.1 Understand the key contributing factors for users to give high stars from LR coefficients

From this example, we observed that

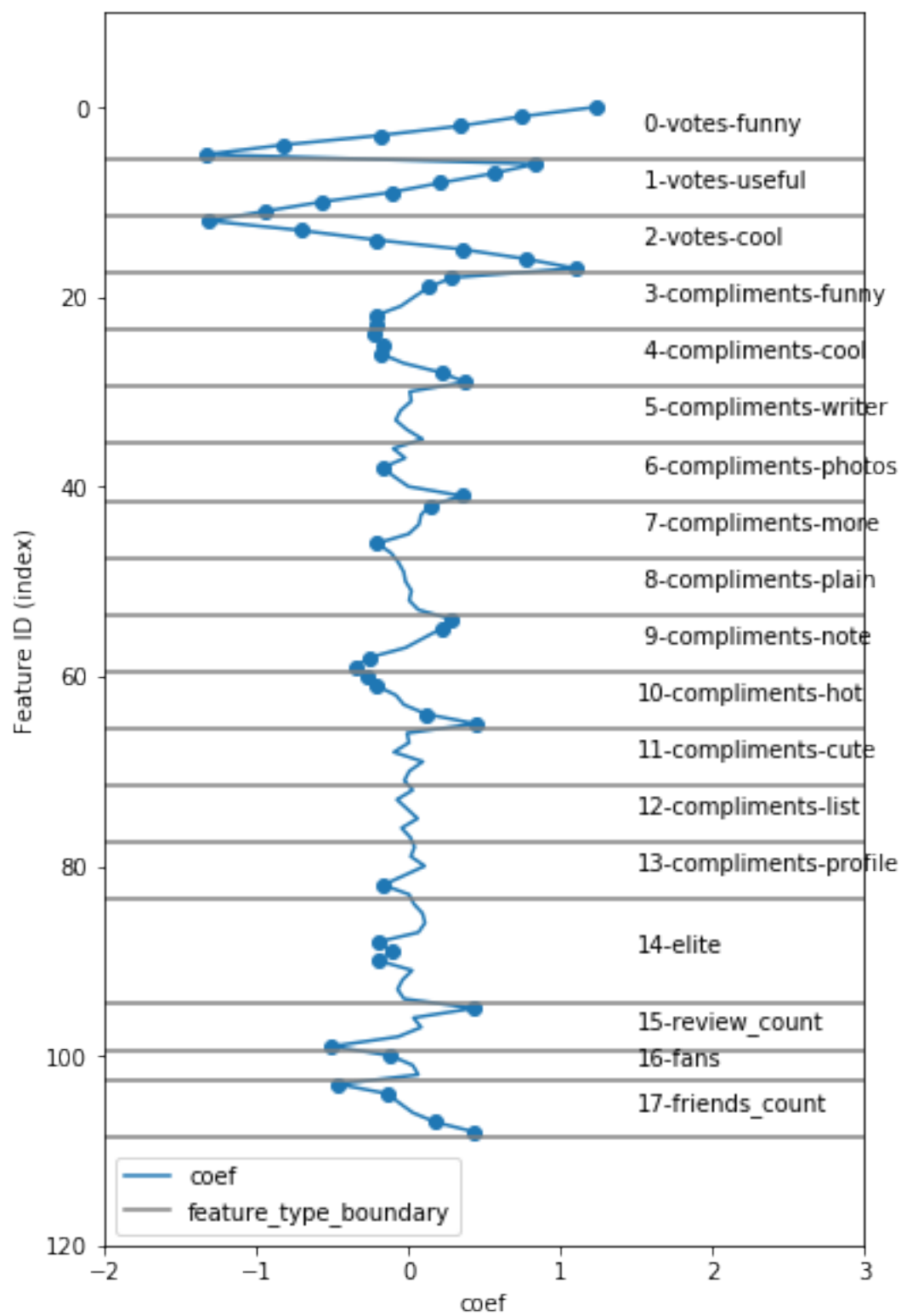
"votes" plays the most dramatic role in how likely an user give high/low-star reviews; more votes on "funny" (also more votes on "useful") lead to a tendency to give low stars, that's probably because the other users feel this user's review more useful when it's negative-trending; in comparison, more votes on "cool" lead to increasing tendency to give high-star reviews

this observations on "votes" aligns well with "compliments"

In addition, increasing review\_counts are negatively correlated with high-star reviewers, and increasing friends (or fans) are positively correlated with high-star reviewers.

Some other features play less significant role, such as "elite" (elite\_year), and "compliments-cute"

```
In [2]: train.plot_coefficients(feature_names, coef, coef_topK, fig_path_coef, False)
```



## 2.2 Table display of top-coefficient features

In [3]: coef\_topK\_df

```
Out [3]:
```

	index	coef	feature	abs(coef)
0	5	-1.333009	votes-funny_bucket_5	1.333009
1	12	-1.314414	votes-cool_bucket_0	1.314414
2	0	1.236864	votes-funny_bucket_0	1.236864
3	17	1.099358	votes-cool_bucket_5	1.099358
4	11	-0.947707	votes-useful_bucket_5	0.947707
5	6	0.840397	votes-useful_bucket_0	0.840397
6	4	-0.825949	votes-funny_bucket_4	0.825949
7	16	0.768446	votes-cool_bucket_4	0.768446
8	1	0.747832	votes-funny_bucket_1	0.747832
9	13	-0.709296	votes-cool_bucket_1	0.709296
10	10	-0.567951	votes-useful_bucket_4	0.567951
11	7	0.558934	votes-useful_bucket_1	0.558934
12	99	-0.502507	review_count_bucket_4	0.502507
13	103	-0.460934	friends_count_bucket_1	0.460934
14	65	0.442499	compliments-hot_bucket_5	0.442499
15	95	0.430626	review_count_bucket_0	0.430626
16	108	0.427941	friends_count_bucket_6	0.427941
17	29	0.374659	compliments-cool_bucket_5	0.374659
18	41	0.356768	compliments-photos_bucket_5	0.356768
19	15	0.352897	votes-cool_bucket_3	0.352897
20	59	-0.351376	compliments-note_bucket_5	0.351376
21	2	0.343260	votes-funny_bucket_2	0.343260
22	54	0.286827	compliments-note_bucket_0	0.286827
23	18	0.284458	compliments-funny_bucket_0	0.284458
24	60	-0.264442	compliments-hot_bucket_0	0.264442
25	58	-0.250540	compliments-note_bucket_4	0.250540
26	24	-0.228637	compliments-cool_bucket_0	0.228637
27	28	0.217751	compliments-cool_bucket_4	0.217751
28	22	-0.217714	compliments-funny_bucket_4	0.217714
29	55	0.216078	compliments-note_bucket_1	0.216078
30	14	-0.215957	votes-cool_bucket_2	0.215957
31	23	-0.212488	compliments-funny_bucket_5	0.212488
32	8	0.210940	votes-useful_bucket_2	0.210940
33	61	-0.208184	compliments-hot_bucket_1	0.208184
34	46	-0.206099	compliments-more_bucket_4	0.206099
35	88	-0.198306	elite_bucket_4	0.198306
36	90	-0.190485	elite_bucket_6	0.190485
37	3	-0.187964	votes-funny_bucket_3	0.187964
38	26	-0.180588	compliments-cool_bucket_2	0.180588
39	107	0.178420	friends_count_bucket_5	0.178420
40	25	-0.171264	compliments-cool_bucket_1	0.171264
41	38	-0.167928	compliments-photos_bucket_2	0.167928
42	82	-0.164595	compliments-profile_bucket_4	0.164595
43	42	0.142703	compliments-more_bucket_0	0.142703

44	104	-0.137408	friends_count_bucket_2	0.137408
45	19	0.135746	compliments-funny_bucket_1	0.135746
46	64	0.120751	compliments-hot_bucket_4	0.120751
47	100	-0.114375	fans_bucket_0	0.114375
48	9	-0.113579	votes-useful_bucket_3	0.113579
49	89	-0.113235	elite_bucket_5	0.113235