



Near Truth Detection

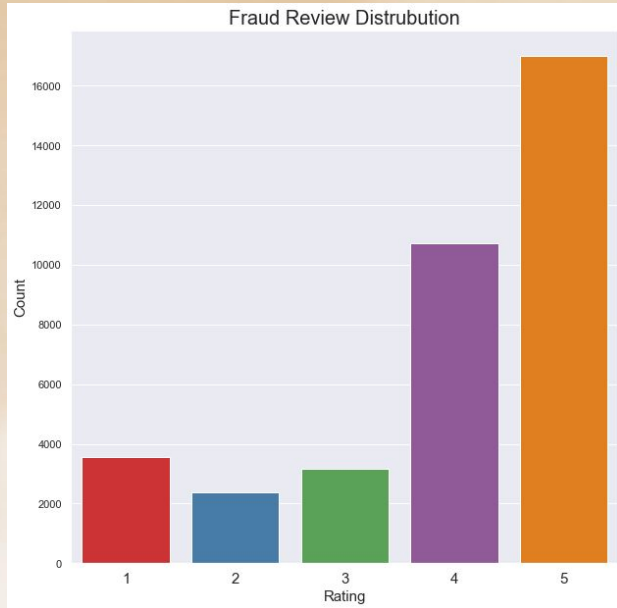
Eddie Lin

Agenda

- Importance
- Data
- Modeling
- Demo
- Next Steps



Why study Reviews?



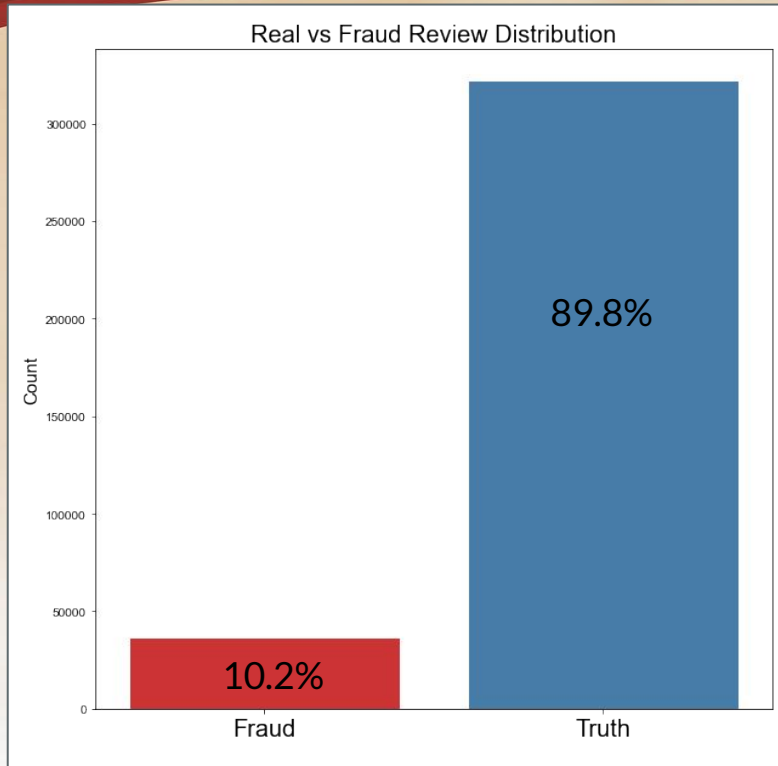
- 2006: 5% → 2013: 20%*
- Additional star improves 5 - 9%**
- 75% people trust positive reviews***
- 60% deterred from negative reviews

*<https://www.businessinsider.com/20-percent-of-yelp-reviews-fake-2013-9>

**<https://www.hbs.edu/faculty/Pages/item.aspx?num=41233>

***<https://blog.reputationx.com/online-reputation-management-statistics>

Data



- 923 Restaurants
- 160,225 Users
- 359,957 Reviews

Source: <http://odds.cs.stonybrook.edu/yelpnyc-dataset/>

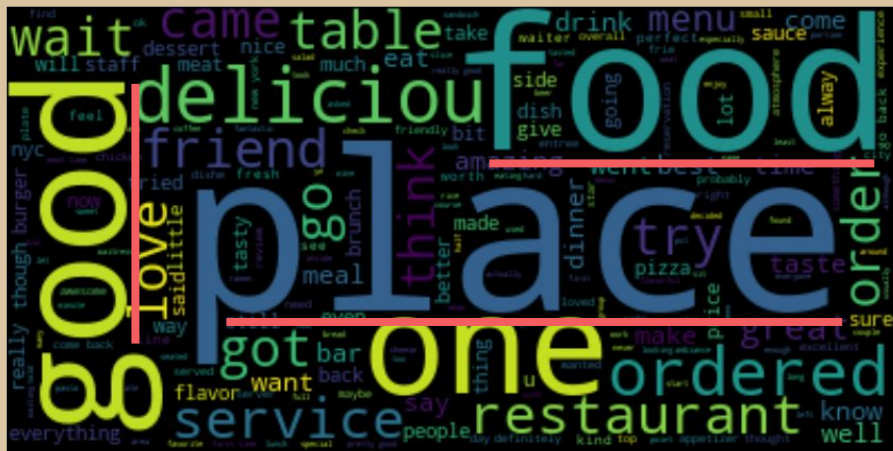


TRUTH



Data

TRUTH



UNTRUTHFUL



Data

- Punctuations and numbers removed
- Short reviews were not considered

“yummy!!1”
“yummy!!!!”
“yummy!”
“yummy.” } “yummy”

“!”

“_”

“.”

“zomg yes”

“yummy place!!!”

“yummy food”

“yummy burger”

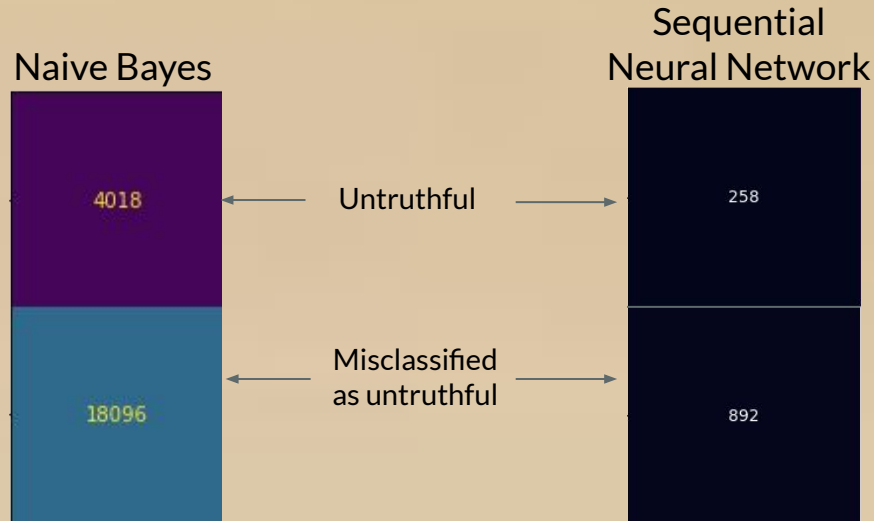


Models

Models	Train	Test	Unseen Data
Naive Bayes	68.57%	70.77%	87.47%
Logistic (w/ features)	52.72%	53.55%	N/A
Random Forest (w/ features)	60.52%	61.07%	N/A
Naive Bayes (w/features)	52.21%	53.01%	N/A
Sequential NN	90.47%	89.02%	88.94%
Sequential NN (w/ embedding)	89.83%	89.80%	N/A

Models (Unseen Data Performance)

- 71,499 unseen
- Naive Bayes: 30% attempt prediction
- Sequential: 8% attempt prediction



Demo

Review From Yelp

Came by for lunch before going to my doctor's appointment as it is on the way. There was a bit of a wait as they were cleaning & sanitizing for our safety (which is great). Each guest is given a large zip lock baggie of napkin, disposable chopstick, plastic soup spoon and paper menu.

I haven't been to a Ichiran in a long time so I had to look over and saw new items (or maybe I forgot) and that it was only tonkotsu ramen available. I vaguely remember other types before but it's fine. I love ramen so I was okay with it. I chose extra toppings to try. Water is self serve (water cooler was right behind me).

Tonkotsu ramen, noodles medium, rich medium broth with 1 garlic clove and extra spicy.
Recommended topping set has soft boiled egg, seaweed, mushroom, 2 extra slice chashu
1/2 order Kaedama (Extra noodles- firm)
Premium Yakibuta

The premium yakibuta was succulent. Juicy, soft tender and flavorful mMmMm. Yum. I inhaled it and requested the kaedama halfway through slurping my ramen so that it would be ready. There was still a wait, I left my bowl alone to save the broth, egg and yakibuta for the extra noodles.

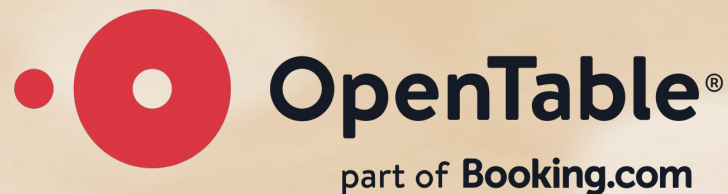
Very slurpious! Yum.

Fake Review From Dataset

the lack of informative details caused a terrible incident to a fellow friend that led to a hospitlization. Any small detail not given of a certain dessert or even food can lead to any harm ..

Next Steps

- More user info
- Compare reviews
- Pass through a link



The Google logo, consisting of the word "Google" in its signature multi-colored font (blue, red, yellow, green, red).



THANK YOU!



Eddie Lin

eddie.jb.lin@gmail.com

Github: RedDragonfruit

Linkedin: eddie-jb-lin

Repo QR Code



https://github.com/RedDragonfruit/Yelp_Near_Truth_Analysis