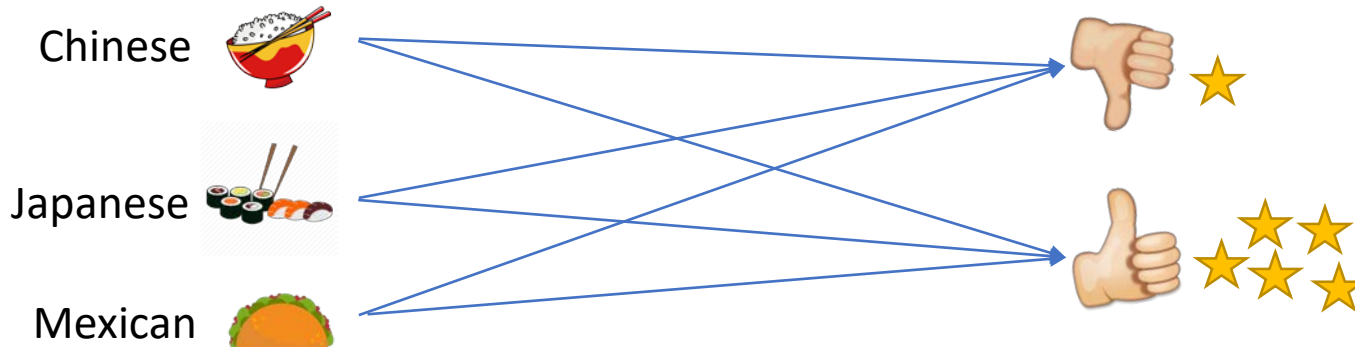# Generating Fake Yelp Restaurant Reviews with RNN

Bowen Li & Sui Huang

Goal: For a specific kind of restaurant, generate good/bad(different star) fake reviews

- **Method:**

1. Train a **_general model(good/bad review)_** then **_customize the food name_**

2. Train a **_"customized" model_**(6 we got) to generate customized reviews

Train a general model then customize the output

Chinese 1 star — Top words for 1 star Chinese restaurant

Japanese 1 star — Top words for 1 star Japanese restaurant

Mexican 1 star — Top words for 1 star Mexican restaurant

Chinese 5 star — Top words for 5 star Chinese restaurant

Japanese 5 star — Top words for 5 star Japanese restaurant

Mexican 5 star — Top words for 5 star Mexican restaurant

Train a general model then customize the output
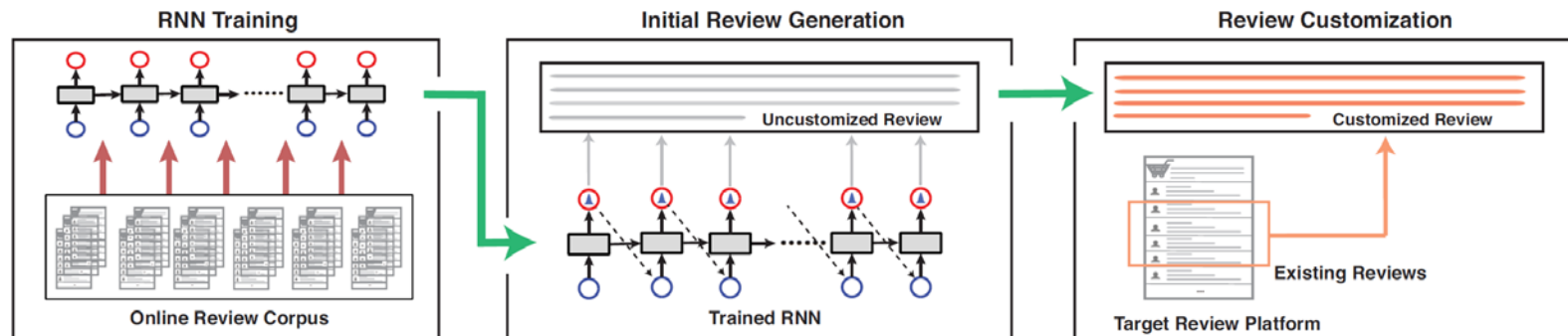
Data Preprocessing
- Only keep English words
- Mark starts and ends of reviews

Generating Initial Reviews
- Based on the size of dataset, adjust the number of neurons (2 hidden layers, each with 1,024 hidden units)
- Train the model

Review Customization
- Use WordNet to calculate lexical similarity
- High similarity -> replace! (food -> taco, sushi)

RNN Training — Online Review Corpus

Initial Review Generation — Uncustomized Review — Trained RNN

Review Customization — Customized Review — Existing Reviews — Target Review Platform

Train a general model then customize the output

"This place is ok. The guy had the grill too high and slightly burnt most of our food. Sushi looked pretty good, but I didn't try it. We'll be sticking with Ah So as their food is much better tasting""說實在話，因為在yelp的評價不低，所以來嚐嚐看，也剛好認識的韓國朋友推薦sashimi don,所以菜單都沒看就點了，服務生很熱情！不過剛端上來時聞到一股濃濃的麻油味我就想說糟糕了，很懷疑為何這樣的don為何需要麻油，夾了魚肉的第一口就吃到不是太新鮮的味道，不過也不至於是無法入口的味道，總之麻油味道把他蓋了過去，後面的幾口不好的味道就慢慢的沒感覺到了，不知道是痲痹了還是剛好第一口夾到地雷。整體來說，這樣的價位也差不多這樣的品質吧！""bomb ass food, horrible service though""bomb ass food, horrible service though""The food is good, but not worth spending $100 per person. I've had better sushi in town for cheaper price"

"<SOR>This place is ok. The guy had the grill too high and slightly burnt most of our food. Sushi looked pretty good, but I didn't try it. We'll be sticking with Ah So as their food is much better tasting.<EOR>"
"<SOR>bomb ass food, horrible service though<EOR>"
"<SOR>The food is good, but not worth spending $100 per person. I've had better sushi in town for cheaper price.<EOR>"

Train a general model then customize the output

Chinese

- This **tofu** is amazing! The food was delicious!!! I was there for a game and on our **dumpling** break for 12. 99 and I love it. For **noodle** was also excellent.

- **dumpling** poutine! Been going to Zo **tofu** for years. We always have a great time there.

Japanese

- The staff were very friendly and attentive. Very fast and friendly service and great **tuna**.

- This **sushi** has some of the best tasting in west **sushi**. It's a **tuna** and fun the next day of going. Very busy but nothing comes close to **sushi** in the kitchen.
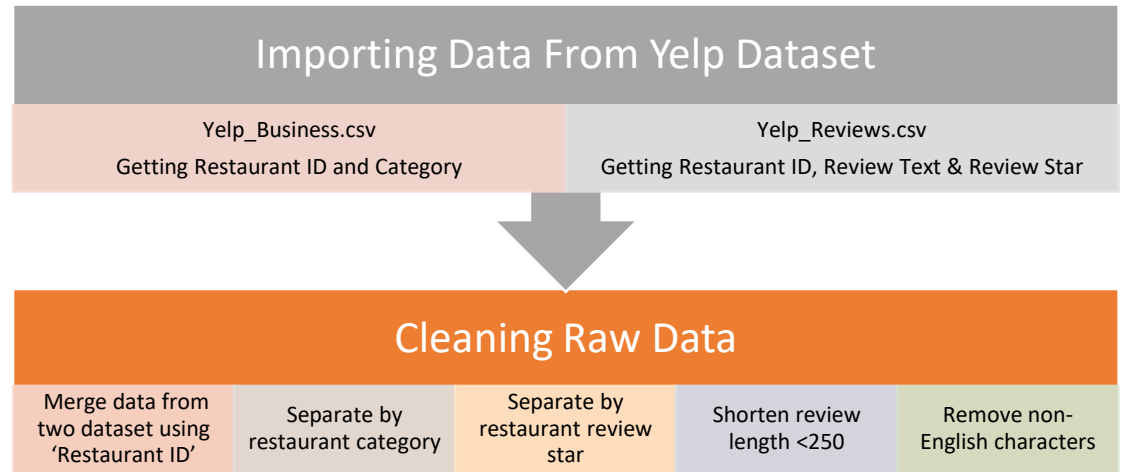
Mexican

- **taco** food and great price! Something different from the convenience of the **taco** but homemade **guacamole** food is pretty good.

- Loved this **taco**. Loved it. Service was excellent, price was right and the atmosphere was kinda high! Can't wait to go back! Good **taco** to dine with friends!

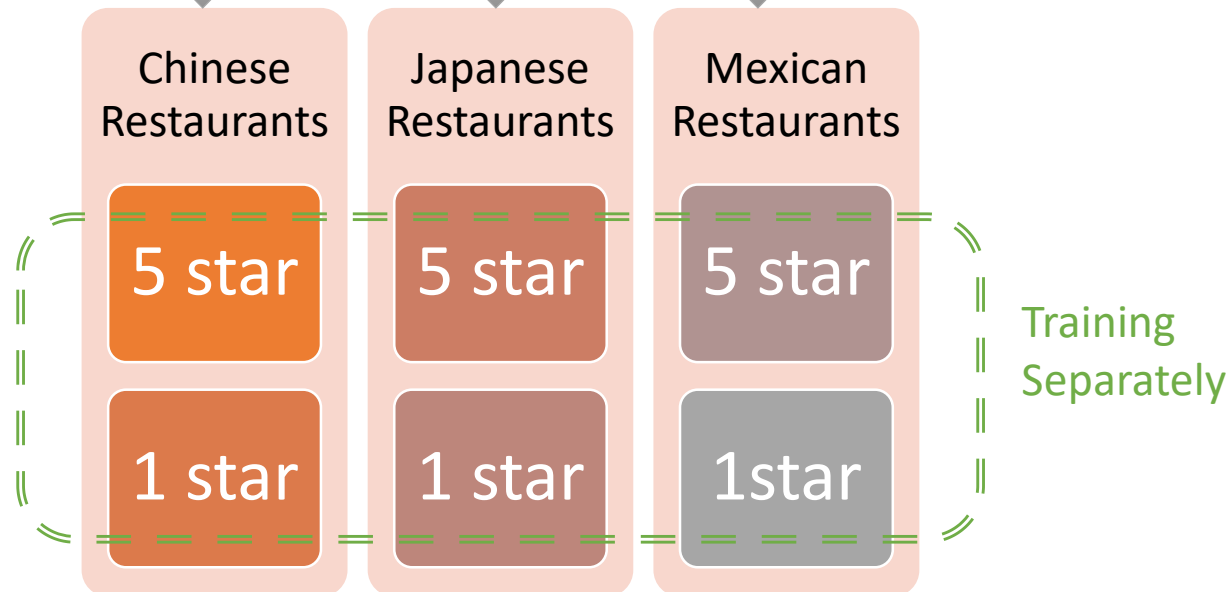# Train a "customized" model to generate customized reviews

Different model for different kinds of restaurant with different review stars.

business_id	name	categories	stars	text
0	PfOCPjBriQAnz__NXj9h_w	"Brick House Tavern + Tap"	American (New);Nightlife;Bars;Sandwiches;Ameri...	4	I've always enjoyed my time at brick house foo...
1	PfOCPjBriQAnz__NXj9h_w	"Brick House Tavern + Tap"	American (New);Nightlife;Bars;Sandwiches;Ameri...	2	1st time here. Came w my Unc bc Louies was clo...
2	PfOCPjBriQAnz__NXj9h_w	"Brick House Tavern + Tap"	American (New);Nightlife;Bars;Sandwiches;Ameri...	1	Worse service ever andI use to be a server so...
3	PfOCPjBriQAnz__NXj9h_w	"Brick House Tavern + Tap"	American (New);Nightlife;Bars;Sandwiches;Ameri...	5	I am updating my review to 5-stars because I t...
4	PfOCPjBriQAnz__NXj9h_w	"Brick House Tavern + Tap"	American (New);Nightlife;Bars;Sandwiches;Ameri...	4	I enjoyed this place. I went the night the Bu...

Importing Data From Yelp Dataset

Yelp_Business.csv	Yelp_Reviews.csv
Getting Restaurant ID and Category	Getting Restaurant ID, Review Text & Review Star

Cleaning Raw Data

Merge data from two dataset using 'Restaurant ID'	Separate by restaurant category	Separate by restaurant review star	Shorten review length <250	Remove non-English characters

Chinese Restaurants | **Japanese Restaurants** | **Mexican Restaurants**

5 star | 5 star | 5 star

1 star | 1 star | 1star

Training Separately

A smaller model

```
Layer (type)                 Output Shape              Param #
=================================================================
lstm_1 (LSTM)                (None, 60, 512)           1722368
_____
lstm_2 (LSTM)                (None, 512)               2099200
_____
dense_1 (Dense)              (None, 328)               168264
=================================================================
Total params: 3,989,832
Trainable params: 3,989,832
Non-trainable params: 0
```

Train a "customized" model to generate customized reviews

Input = 'This is a text file. This file is just something to help the model get started! You can use anything.'

Chinese

- **(5 star)** This is my new favorite Chinese place in Vegas. The servers are friendly. We had the chicken pakora, chicken wings, and shrimp with broccoli is AMAZING too. We love it!
- **(1 star)** The food was terrible. Worst chinese food ever!!! The service was really bad, the food was horrible. I wouldn't recommend this place to anyone.

Japanese

- **(5 star)** This is a great sushi place in the world! Sushi is amazing, the price is reasonable and the prices are perfect. The staff is super friendly and the service was very even better. Happy hour selection was great and the food is fantastic. I will definitely be back.
- **(1 star)** Poor attempt at Japanese noodle .The food is decent, but I'd rather give my order. Food was mediocre at best.

Mexican

- **(5 star)** The best Mexican food in Vegas. The food is amazing and the salsa bar is very good.
- **(1 star)** Sat down excited to eat here. Been there into top mun right mexican food. Over priced for what you get. Will not be returning to this place anymore!

Comparison

Train a general model then customize the output

Universal

Once the training is finished, all it needs is customization.

Might end up with weird reviews that can be easily recognized.

Using lots of resource to train.

Different model for different kinds of restaurant with different review stars.

Specific

Each kind of restaurant gets their own models for good/bad reviews.

Limited data, unusual reviews(emoji, 中文), the model might not converge.

Defense
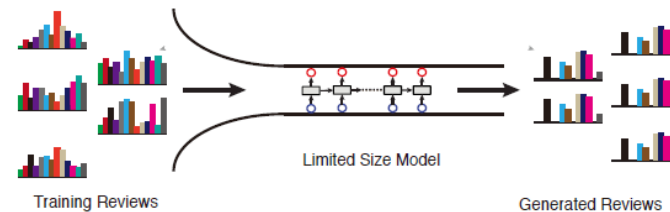
ML-based Review Filter(Classifier):

Does not work.

- A supervised ML scheme based on **linguistic** features (feature filter)

- A plagiarism detector to check for **duplications**

- Human User Study

RNNs-Based Defend:

Looking for probability.

This leverages the fact that a generative language model **builds a fixed memory representation** of the entire training corpus, which limits the amount of information that can be learned from a training corpus

Training Reviews   Limited Size Model   Generated Reviews

Algorithm 1 Proposed Defense

▷ input-R_F:machine-generated review training set, R_L:real review training set, T:test review
1: **procedure** DEFENSE(R_F, R_L, T)
2:     N ← length(T)
3:     RNN_F ← trainRNN(R_F)
4:     RNN_L ← trainRNN(R_L)
5:     **for** t = 1:N-1 **do**
6:         feed X_t into RNN_F
7:         $\mathcal{L}_F$ ← $P_F(X_{t+1}=x_{t+1}|x_1,...,t)$
8:         feed X_t into RNN_L
9:         $\mathcal{L}_L$ ← $P_L(X_{t+1}=x_{t+1}|x_1,...,t)$
10:        $\mathcal{L}_t$ ← $-\log \frac{\mathcal{L}_L}{\mathcal{L}_F}$ ▷ negative log-likelihood ratio
11:    $\bar{\mathcal{L}}$ ← $\frac{\sum_{j=1}^{N-1} \mathcal{L}_i}{N-1}$
12:    **if** $\bar{\mathcal{L}} > 0$ **then**
13:        return FAKE
14:    **else**
15:        return REAL

Train two RNNs, based on real and fake reviews. Test after which RNN, the input would have a **high probability** of generating next letter in this RNN.

Yao, Yuanshun, et al. "Automated Crowdturfing Attacks and Defenses in Online Review Systems." Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. ACM, 2017.