W266 Final Project

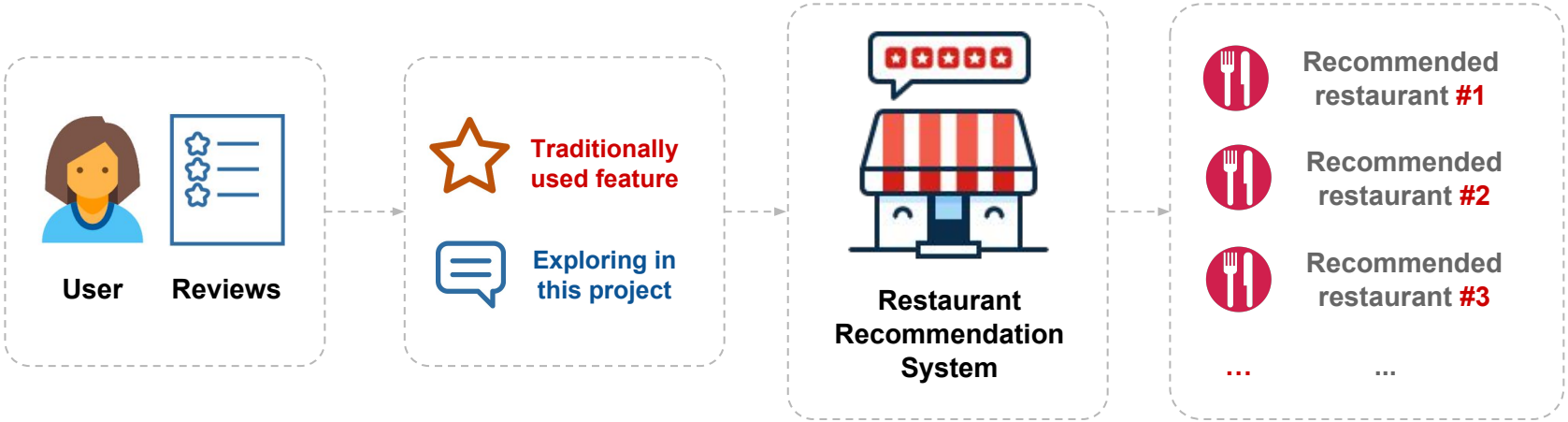
Restaurant Recommendation System Based on Yelp Data

Aaron Yuen, Alexander Herring, Charlene Chen

8/7/2018

Problem Statement

- We explored using a recurrent neural network (RNN) with LSTM to generate a recommendation system for Yelp restaurants, primary leveraging Yelp restaurant review text.
- We also explored an **experimental** topic modeling approach (collaborative filtering with LDA) for generating recommendations without using neural networks or word embeddings.



Dataset

The raw Yelp dataset consists of:

- 2.6M business reviews
  user id, review text, star rating, and cool, funny, useful scores
- 51k businesses
  name, business categories, city, lat/long
- 367k users
  review count, Yelping since, average stars, etc.

After processing and filtering:

- Non-restaurant businesses filtered out
- Users with 5 or less reviews filtered out
- Final dataset: **1.5M reviews for training, 373k for testing**

* Yelp dataset (https://www.yelp.com/dataset)

Models - Overview

Three approaches were used for this project:

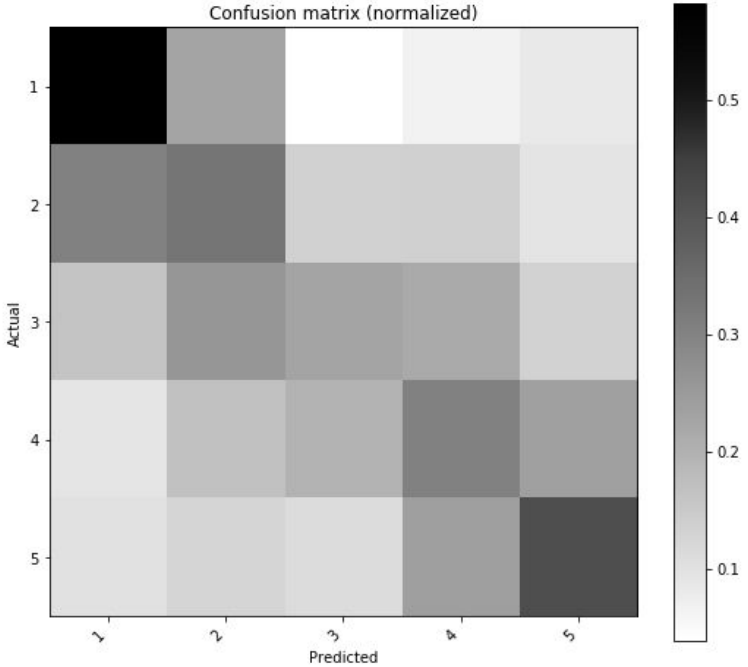
Approach	Short Description
Gaussian Naive Bayes (Baseline)	• A Naive Bayes model leveraging continuous GloVe word embeddings • Generates sentiment analysis for reviews by predicting star ratings
Recurrent Neural Network (RNN) Model with LSTM	• A Recurrent Neural Network (RNN) model leveraging continuous GloVe word embeddings • Generates sentiment analysis for reviews by predicting star ratings
Experimental - Collaborative Filtering Using Topic Modeling (LDA)	• Leveraging review text to derive a set of topics that represents the review corpus • With the trained LDA model, a user-topic distribution matrix is built • For a given user, highly-rated restaurants from similar users in the user-topic-distribution space will be recommended

Evaluation - Accuracy & F1 Scores

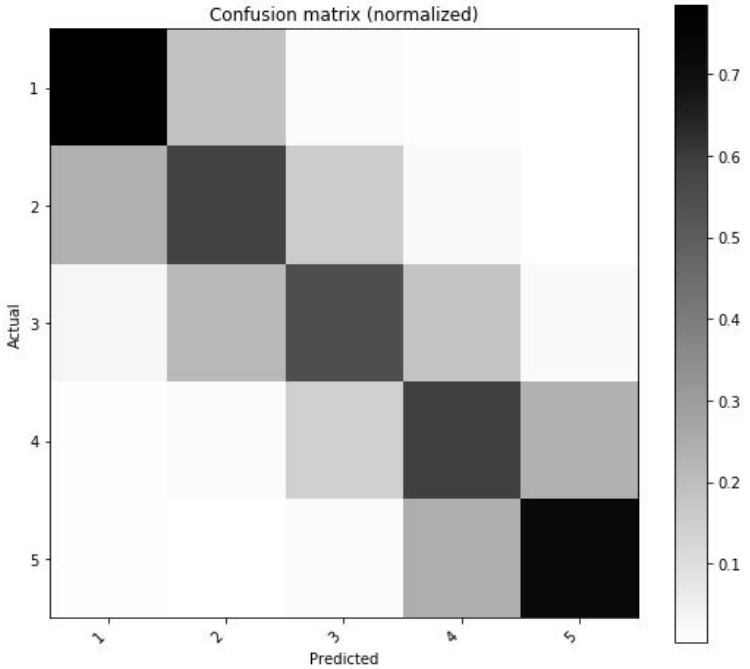
Model	Accuracy	F1 Score
Gaussian NB (Baseline)	35.1%	0.36
RNN-LSTM	**64.6%**	**0.65**
Collaborative Filtering w/ LDA Topic Model (LDA) (Experimental)	37.0%**	0.43

** Not a direct comparison given that an “accurate” label requires the CF recommendation system to have recommended a restaurant that the user has already given a review on.

Evaluation - Confusion Matrices



Gaussian NB
(Baseline)



RNN-LSTM

Error Analysis - RNN-LSTM

Most errors in predicted star ratings were within ± 1 stars of the actual ratings:

Example Review:

I can't believe all the negative reviews. I've been eating at Bubbas' for years and the food is always top notch. Best Eastern, NC bbq in Charlotte. Get the large plate and include the Brunswick stew. Sweet tea is strong and tasty,

Predicted Stars: 5

Actual Stars: 4

Error Analysis - RNN-LSTM

Reviews with a spread of more than ± 2 stars often used more positive or negative language than the review seem to warrant:

Example Review:

I had been to this plaza 100s of times in my life and never noticed **this gem of a dive bar & grill**. Good service, food & prices. The Mediterranean Dip is **crazy delicious**. I wish it was socially acceptable to order it and just eat it with a spoon!

Predicted Stars: 5

Actual Stars: 3

Error Analysis - RNN-LSTM

Reviews with a spread of more than ± 3 stars had sections that read as if they had the opposite sentiment of the overall review score (e.g. the reviewer seemed angry or hostile during sections of a positive review)

Example Review:

For anyone trying to compare this buffet to one that is double or triple the price, don't. Just get out your fat wallet and move over the strip. For others of us who want **decent value with good, everyday food**, this is the place. I'm not overly picky and I like non complicated food that I can identify with, I guess you'd call it comfort food. I've been to all those expensive buffets and I still say for the value Palace Station wins.

Predicted Stars: 2

Actual Stars: 5

Error Analysis - RNN-LSTM

Reviews with a spread of ± 4 stars also had sections that seemed to contradict the review score, and often had sarcasm. They also tended to be longer.

Example Review:

Diabetics: **Beware. Stay away.** Buy some for me and **don't eat any! ;)**

OMG, **soo good**, probably the **best butter tarts that I've had** AND I'm not even a sweets person! Definitely sweet though. Wash down with unsweetened coffee/tea.

About $2.50-3ea. 2 tarts were like $6 'n change.

I sorta 'forgot' about the tarts (Plain, Pecan) for a coupla days after SoupFest - long story - but when I ate them, they tasted fresh still. An indication of the amount of sugar perhaps?

Pastry tasted like a butter cookie. **Will never buy that grocery store crap again** as it can't even compare! Mmm...drool...

Tip: A few spots around back of the building, in front of skatepark.

Predicted Stars: 1

Actual Stars: 5

Error Analysis - LDA (Experimental)

Generally, topics derived from the LDA model (100 in total) were hit and misses:

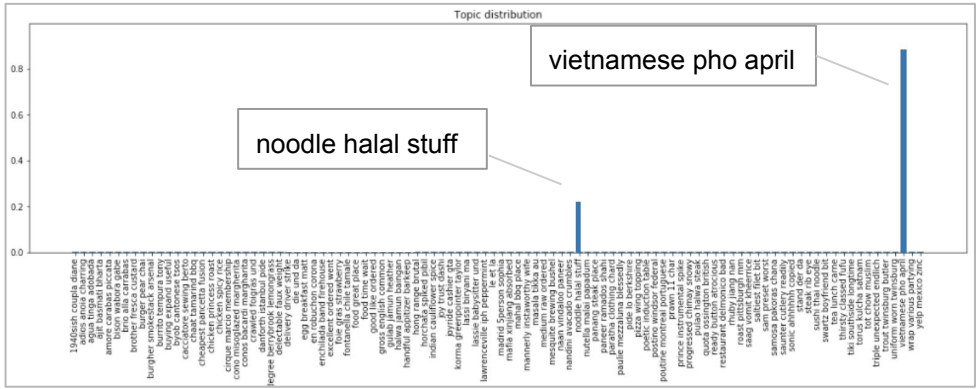
Topic #62: masala tikka
Topic #76: vietnamese pho
Topic #95: steak rib eye
Topic #96: naan vindaloo paneer

Topic #5: food great place
Topic #63: excellent ordered went

Topic #72: uniform worn twinsburg
Topic #89: swartz boyfriend bit

Example of a **good** review-topic distribution:

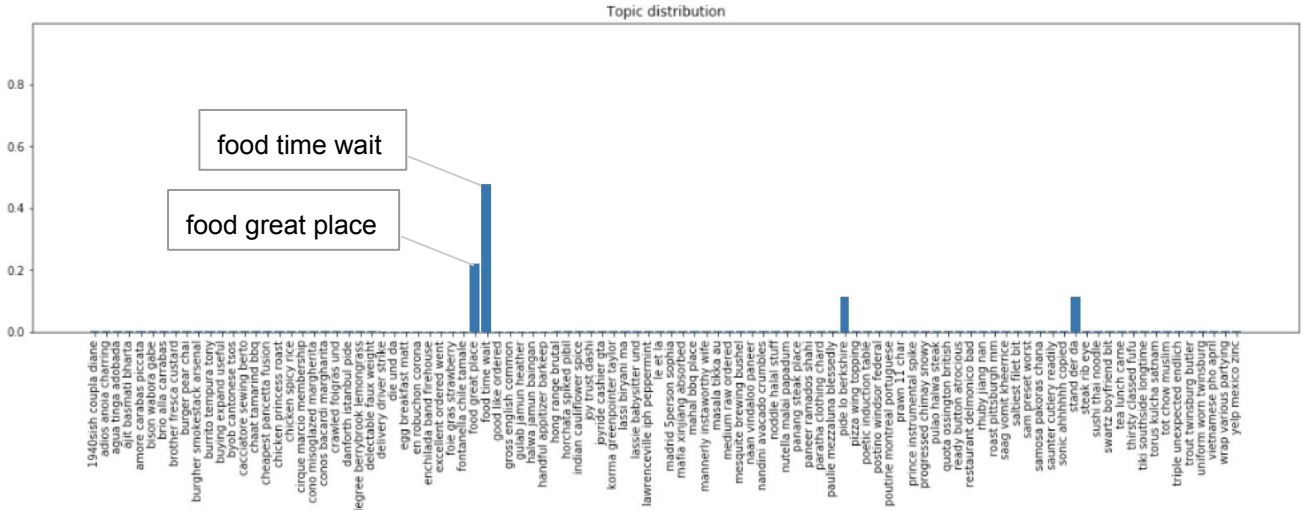
Pho King **Vietnamese** Cuisine has a really strange menu... aside from the usual **Vietnamese Pho**, Bun or whatever... it is the first **Vietnamese** restaurants I been to offering Mala (numbing spicy) lamb noodle soup?! And there are quite a few lamb dishes on the menu too… But coming here as a first timer, it is hard for me not to stay with my usual choice at a **Pho** joint? I will leave those lamb dishes for the more adventurous type!\n\nSpecial House **Pho** Small ($5.75)
The broth was nicely seasoned and with a hint of herbs and anise. Beef Meat was plentiful, the tendon was especially tende. **Noodle** was the fresh type. Size wise the portion is more like a medium in a lot of places!!



Error Analysis - LDA (Experimental)

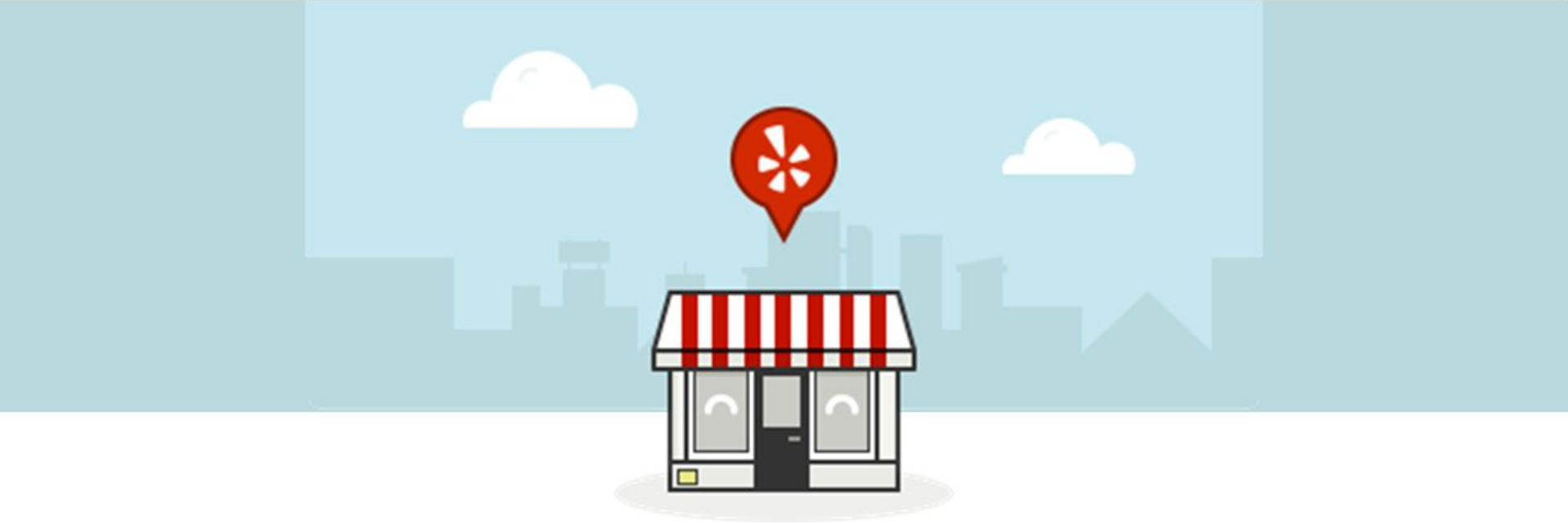
Example of a **bad** review-topic distributions:

i have eaten here a couple **times** and i don't think i was unhappy about anything other than their prices. Its kind of expensive to me but it taste **great**. They have one of the best turkey burgers I've had in a long **time** and their banana pudding is pretty tasty. The **food** is made to order and their cashiers are very friendly. Its nice and clean, plus its located in a nice shopping mall.



To Take it Further

Approach	Next Steps
Recurrent Neural Network (RNN) Model with LSTM	• Training optimizations (current model takes **>21 hrs** to train!) • Experimentation with different word embedding methods than GloVe (e.g. Word2Vec, Doc2Vec)
Collaborative Filtering Using Topic Modeling (LDA)	• Training optimizations (current model takes **>12 hrs** to train!) • Better parameter tuning for LDA model to get more meaningful topics
General	• Incorporate other review information (usefulness, funny counts, etc.) for the recommendation ranking score • Factor latitude/longitude of businesses into recommendations (e.g. user in Seattle should not get recommendations for restaurants in Florida)



Thank you!

Appendix

Recurrent Neural Network Model with LSTM

Collaborative Filtering Using Topic Modeling (LDA)

Latent Dirichlet Allocation (LDA) is a generative statistical model that, when trained, assigns <MORE HERE>