

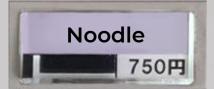
食券をお買い求めくださし

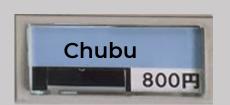
両替を希望されるお客様は、スタッ お申し付けください。



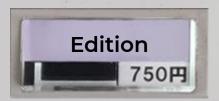














Problem:

- Small shops informal
 - No websites
- Tabelog has data
 - All in Japanese



NLP Workflow and Tools:







Web Scraping



Tokenizing and POS Tagging







NumPy+pandas+scikit-learn

Topic Modeling with Latent Dirichlet Allocation



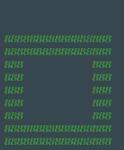
Recommendation Web App

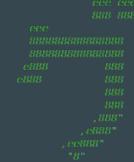




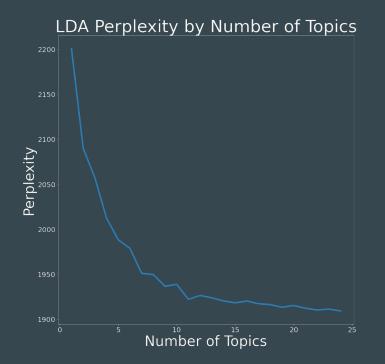








LDA Topic Model:

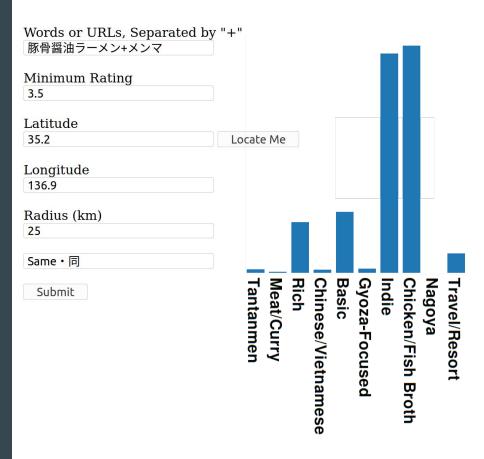


Topic Label	Representative Term	Translation
Tantanmen	担々麺	Tantan noodles
Meat/Curry	ヒレカツ	Fillet cutlet
Rich	フジヤマ55	Fujiyama 55 (a chain)
Chinese/Vietnamese	回鍋肉	Twice-cooked pork
Basic	吉村家	Yoshimuraya (a chain)
Gyoza-Focused	浜松餃子	Hamamatsu gyoza
Indie	赤丸	Akamaru (shop name)
Chicken/Fish Broth	鶏白湯	White chicken broth
Nagoya	名古屋メシ	Nagoya food
Travel/Resort	平湯温泉	Hirayu Hot Spring

Functionality:

- 1. Topic vector of search
- 2. Sort similar or different
- 3. Filter by geography
- 4. Filter by review score

Next Noodle Shop ・ 次の麺屋



Tonkotsu shoyu ramen + bamboo shoots

→ Pretty inoffensive ramen



Next Noodle Shop · 次の麺屋



Different · 别

→ Spicier, More Chinese



Next Noodle Shop · 次の麺屋

Words or URLs, Separated by 豚骨醤油ラーメン+メンマ	/ " + "	
Minimum Rating		
3.5		
Latitude		
35.2	Locate Me	1-4 1
Longitude		plot here
136.9		
Radius (km) 25		
Different • 別		
Submit		

Recommendations:

かっぱ園菜館 中国上海料理 豫園 想吃担担面ユニモール店 スコンター錦店 豫園 金山店

Clear/Reset

かっぱ: Japanese Monster

- Fresh vegetables
- Cold (sometimes)
- Strong flavors

Search+White Chicken Broth

Drier, Stronger Flavors, Even More Chinese



Next Noodle Shop · 次の麺屋

Words or URLs, Separated by 豚骨醤油ラーメン+メンマ+鶏白湯	/ "+" 	
Minimum Rating		
3.5		
Latitude		
35.2	Locate Me	
Longitude 136.9		plot here
Radius (km)		
25		
Different · 別		
Submit		

Recommendations:

中国上海料理豫園かっぱ園菜館スコンター錦店想吃担担面ユニモール店豫園金山店

Clear/Reset

中国上海料理: Shanghai Cuisine

担担面:Tantanmen/DandanMein

- Written Chinese way
- Still strong flavors

Minimum Rating 3.7

→ Chinese Focus Diluted



Next Noodle Shop ・ 次の麺屋

Words or URLs, Separated by 豚骨醤油ラーメン+メンマ+鶏白湯	⁷ "+"	
Minimum Rating		
3.7		
Latitude		
35.2	Locate Me	
Longitude 136.9		plot here
Radius (km) 25		
Different · 別		
Submit		

Recommendations:

味仙矢場店 味仙今池本店 土夢 つけ麺丸和春田本店 からみそラーメンふくろう本店

Clear/Reset

本店: Original Location

- ¾ honten (本店)
- 2 same chain
- Still strong flavors

Minimum Rating 3.7

→ No. 5 still strong and spicy



Next Noodle Shop · 次の麺屋

Words or URLs, Separated by 豚骨醤油ラーメン+メンマ+鶏白湯	"+"	
Minimum Rating		
3.7		
Latitude		
35.2	Locate Me	
Longitude 136.9		plot here
Radius (km) 25		
Different · 別		
Submit		

Recommendations:

味仙矢場店 味仙今池本店 土夢 つけ麺丸和春田本店 からみそラーメンふくろう本店

Clear/Reset

辛味噌: Spicy Red Miso

- Thick noodles
- Fatty broth
- Spicy/strong

Future Work:

- More of Japan
- Translation
- Update search vector plot
- Aesthetic Design



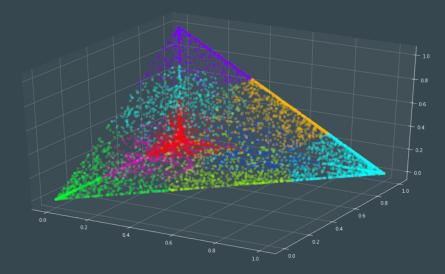
Thank you!

Sam Mize

Linkedin
Email
GitHub



Appendix:



This plot, which was made when LDA was completed with 4 topics, shows K-means clusters plotted in three of the 4 topic dimensions. I was not convinced that there were actually any legitimate clusters and chose to pursue recommendation through topic vectors without clustering instead.

Appendix:

	Tantanmen	Meat/Curry	Rich	Chinese/Vietnamese	Basic	Gyoza-Focused	Indie	Chicken/Fish Broth	Nagoya	Travel/Resort
0	担	とん汁	丸和	回鍋肉	吉村家	浜松餃子	大喜	レアチャーシュー	驛	高山
1	担担麺	カツ丼	カラメ	皿うどん	信州味噌	ハルピン	白丸	鶏白湯ラーメン	オッチャホイ	青島
2	芝麻醬	PA	ラーメン二郎	ニラレバ	麦味噌	8番ら一めん	志太	汐	大名古屋ビルヂング	長岡
3	高麗人参	ヒレカツ	ブッチャー	国士無双	西山製麺	テンホウ	第一旭	オマール	敦賀ラーメン	ヒグマ
4	マーラー	売店	嘉六	あんかけ焼きそば	横浜家系	ハルピンラーメン	赤丸	鶏白湯	錦三	高山ラーメン
5	錦城	八ヶ岳	らいこう	広東	家系ラーメン	くるまやラーメン	大石	たぶつ	今池	安養寺
6	陳麻婆豆腐	下り線	勝龍	中華飯	吉田家	丸源ラーメン	ラーショ	塩そ	まじん	長岡生姜醤油ラーメン
7	彰貴	黒部ダムカレー	やもり	ゲソカラ	六家	べとん	藤	ニボニボ	豚神	長岡市内
8	高浜	黒部ダム	フジヤマ	ホイコーロー	田所	有楽	カナキン	福座	呑助	甚五郎
9	坦	山賊焼	しゃが	かた焼	岩本	人力車	富山ブラック	コーチン	重油	飛騨高山
10	美合	さざえ	安福亭	八宝菜	キャベチャー	下諏訪	久留米ラーメン	セメント	名古屋メシ	小千谷
11	ユニモール	談合坂	三郎	スッポン	濱家	むつ	久留米	鶏ムネ肉	蜃気楼	佐蔵
12	山之手	カレーパン	ら・けいこ	ぼんち	黒田	赤餃子	一番星	穂先	ももちゃん	左近
13	龍龍	かつ丼	ジロー	マーボー丼	山岡家	宝龍	チャーシューワンタンメン	飛魚	酢鶏	宮内
14	城北	釜めし	マシライス	レバニラ炒め	龍上海	ら一めん世界	南京	平子	神風	安養寺ラーメン

Acknowledgements:

Data: www.tabelog.com Images: wikipedia, tabelog