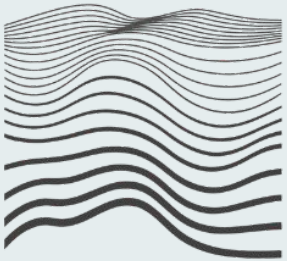
Natural Language Processing

Restaurant Classification

project by:



LINDSAY MCFARLANE

www.lindsaymsba.com

831.601.7684 | lindsay.alexandra14@gmail.com | linkedin: lindsay-mcfarlane

1 Summary

2 Full Deck

Summary

What is the project?

Goal

Build an **NLP model** to accurately **classify restaurant types** with Yelp data

Objective

Outperform given baseline model

What is the data?

10 Classes

Restaurant Types


Mediterranean


Italian


Canadian (new)


Thai


Japanese


American (new)


American (traditional)


Chinese


Asian Fusion


Mexican

61 Features

Restaurant Features


Good for Kids


Price Range


State

Name


Attire


Review


Outdoor Seating


Delivery


Alcohol


Ambience

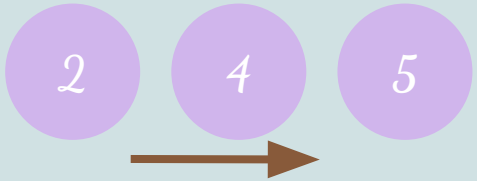
Sample Features

Models

1. Avg. Word Embeddings of **Restaurant Reviews** in Logistic Regression
2. Avg. Word Embeddings of **Reviews + other Features** in Logistic Regression
3. BERT Transformer Model on **Restaurant Reviews**
4. BERT Transformer Model on **Reviews + other Features**
5. roBERTa Transformer Model on **Reviews + other Features**

Models: Adding Other Features

For models with additional features:



Appended additional features as text to the reviews feature

Review:

So, we stopped here on our way to the Side Quest, which is just up the street. We did not know what to expect, but came on a 3$ taco night...I'm so glad we did:) Overall wonderful experience!

Review + **Other Features:**

Casual. Restaurant named Rush Inn. Moderate Price. Noisy. Located in Ohio. Full bar. Has TV. Caters. Takeout. Casual attire. Good for lunch, Good for dinner. So, we stopped here on our way to the Side Quest, which is just up the street. We did not know what to expect, but came on a 3$ taco night...I'm so glad we did:) Overall wonderful experience!

Results

Baseline:

Top:

Metric	Model 1: Avg. Embeddings Review Only	Model 2: Avg. Embeddings Review+Features	Model 3: BERT Review Only	Model 4: BERT Review + Features	Model 5: RoBERTa Review + Features
accuracy	77.7%	77.7%	80.7%	82.7%	82.1%
macro avg	65.0%	65.4%	75.7%	79.5%	79.0%
weighted avg	75.0%	75.1%	80.3%	82.9%	82.3%

Class	Model 1: Avg. Embeddings Review Only	Model 2: Avg. Embeddings Review+Features	Model 3: BERT Review Only	Model 4: BERT Review + Features	Model 5: RoBERTa Review + Features
american (new)	40.7%	40.4%	48.5%	61.2%	58.7%
american (traditional)	72.2%	72.4%	71.5%	73.8%	72.1%
asian fusion	5.4%	5.4%	48.6%	52.2%	50.7%
canadian (new)	7.1%	10.5%	38.6%	55.7%	59.0%
chinese	85.6%	84.7%	93.2%	91.4%	92.3%
italian	88.4%	88.6%	91.0%	92.1%	90.9%
japanese	86.0%	84.9%	91.7%	89.4%	89.5%
mediterranean	81.7%	82.5%	84.7%	87.9%	85.7%
mexican	95.4%	95.7%	95.9%	96.3%	96.6%
thai	87.9%	88.6%	93.8%	94.5%	94.6%

Results

Model 4 (BERT w/ added features) performed the best:

*achieved **~83%** weighted F1,*

*a +10.5% improvement vs. baseline**

***note:** was limited with compute resources; ideally would do more runs and average them*

notebook link:

Technical Setup

Data

Source: Yelp
(given for Spring '25 UCSD NLP Class)

Type: Structured

Features: 61 restaurant features (including restaurant review)

Target: 10 Classes: Restaurant Types (e.g., Italian, Thai, etc.)

Setup

Language: Python

Models: Word2Vec (Skip-gram), BERT, RoBERTa

Compute: NVIDIA A100 GPU in Google Colab

Evaluation Metrics

Classification:

- F1 Score:
  - Balance of Precision & Recall
- Weighted F1 score due to class imbalance
- Time to run model

Click for Full Deck