Restaurant Recommendation System using Natural Language Processing

Name: Soe Zhao Hong ID: 1181101614

Supervisor: Goh Hui Ngo Moderator: Amy Lim Hui Lan FACULTY OF COMPUTING & INFORMATICS

Introduction

This research focuses on how to recommend restaurants based on the restaurants' reviews using various word embeddings techniques.

Besides, enhanced the recommendation system using Sentiment Analysis and Named Entity Recognition.

Understanding the Problem



What is the way to improve recommendation?

02

How to implement restaurants' review into a recommendation system?



How to evaluate the recommendation system?

Objectives

To propose an algorithmic framework for a restaurant recommendation system



To employ Natural Language Processing for the recommendation of restaurants



To perform evaluations on the models to validate the performances of proposed approaches



Project Scope



Build content based recommendation using word embedding models



Recognize the food name and its sentiment from the review

Literature Review

Recommendation

- Content-based
- Collaborative Filtering
- Location-based
- Hybrid

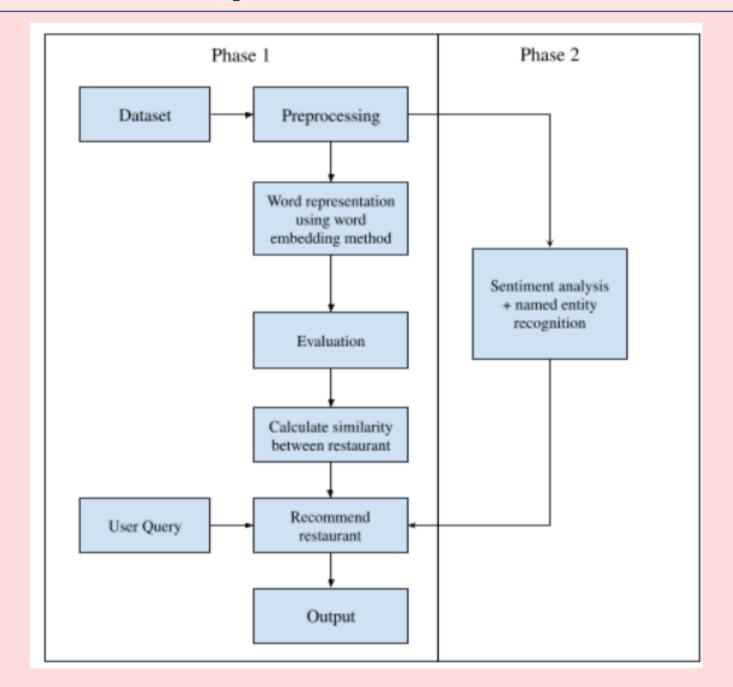
Word Embedding

- TF-IDF
- BERT
- Word2Vec
- CountVectorizer

Sentiment Analysis

- Trained Sentiment Analysis to predict positive and negative sentiment
- Sentiment Analysis help
 to improve the completion
 of the recommendation
 system.

Proposed Framework



Word Embedding

O1 Term Frequency-Inverse Document Frequency

Representations from

Transformers

- Inverse Document
 Frequency
 Bidirectional Encoder
- 04

CountVectorizer

Word2Vec

Evaluation Metrix

Supervised proxy problem (Alexander Geiger, 2021)

- A related supervised learning algorithm to predict the features of the data
- Train supervised learning algorithms: Random Forest, Naive Bayes, and K nearest neighbor(KNN)
- The trained model is used to predict the rating of each testing data after training.

Sentiment Analysis + Named Entity Recognition

- 01
- Pretrained Named Entity Recognition model from DeepNote.com
- 02

Find out the Top-N of Part of Speech with GOOD/BAD sentiment score and the Food entity after using model from DeepNote.com

Result and Discussion

Word Embedding Evaluation Result

Random Forest

Word Embedding Model	Precision	Recall	F1	Accuracy
TF-IDF	0.86	0.82	0.81	0.82
Word2Vec	0.84	0.79	0.78	0.79
BERT	0.87	0.84	0.84	0.84
CountVectorizer	0.83	0.78	0.78	0.78

Naïve Bayes

	,				
Word Embedding Model	Precision	Recall	F1	Accuracy	
TF-IDF	0.89	0.88	0.88	0.88	
Word2Vec	0.99	1.00	0.99	1.00	
BERT	1.00	1.00	1.00	1.00	
CountVectorizer	0.71	0.58	0.59	0.58	

KNN

Word Embedding Model	Precision	Recall	F1	Accuracy
TF-IDF	0.69	0.65	0.66	0.65
Word2Vec	0.71	0.70	0.69	0.70
BERT	0.73	0.72	0.72	0.72
CountVectorizer	0.68	0.65	0.66	0.65

Part of Speech with sentiment and food entity

Positive + Food NER	Negative + Food NER			
'DT', 'NN', 'NN', 'JJ'	'NN', 'PRP', 'VBD', 'JJ', 'NN', 'NN', 'NNS', 'NNP', 'NNS', 'VBD', 'JJ', 'NNS', 'VBD', 'RB', 'JJ', 'NNS', 'VBP', 'NNP', 'NNP'			
'DT', 'NN', 'NN', 'RB', 'JJ'	'NNP', 'JJ', 'NN', 'VBN', 'NN', 'NN', 'NN', 'JJ', 'NN', 'VBG'			
'DT', 'NN', 'JJ', 'NN'	'NNP', 'NNS', 'VB'			
'DT', 'NN', 'NN', 'NN', 'JJ'	'PRP', 'VBP', 'JJ', 'NNS', 'VBP'			
'JJ', 'NN', 'NNS'	'IN', 'NNS', 'JJ', 'NN', 'NNS', 'VBP', 'JJ', 'NN', 'NNS'			

Final Result of Restaurant Recommendation System

cuisines	Restaurant_Name	address	Rating	Service	Recommended	Not Recommended
North Indian, Street Food, Mithai	Laddoos	799, 7th Cross Road, 2nd Stage, Near Fooddays, BTM, Bangalore	3.4	Delivery	'samosas', 'kachoiri'	-
Biryani, Hyderabadi, Chinese	Biryani Palace	L139, 15th Cross, 5th Main, Sector 6, HSR Layout, HSR, Bangalore	3.9	Delivery		
Biryani, Andhra, North Indian, Chinese	Shanmukha	6th Cross Road, HAL 2nd Stage, Old Airport Road, Bangalore	3.8	Dine-out		-
North Indian	Kataria's Pakwan	31, 2nd Cross, 1st Stage, Padmanabh Nagar, Kumaraswamy Layout, Bangalore	3.3	Dine-out	-	-
Asian, Chinese, Continental, Italian	Thyme & Whisk	56/1, Sri Krishna Towers, 9th Main Road, 5th Block, Jayanagar, Bangalore	3.9	Delivery	'broccoli', 'almond soup'	-
American, North Indian, Chinese	The Barn - Bar & Kitchen	5th Floor, Garuda Mall, Magrath Road, Ashok Nagar, Brigade Road, Bangalore	4.2	Pubs and bars	'frnds'	-
Bakery, Desserts, Beverages	Mad Over Donuts	Food Court, Garuda Mall, Magrath Road, Ashok Nagar, Brigade Road, Bangalore	3.9	Delivery	-	-
Continental, Chinese, Fast Food	Amber Rush Restobar	1st & 2nd Floor, 17/4, Residency Road, Bangalore	4.1	Pubs and bars		-
Cafe, Chinese, Pizza, North Indian, Burger	Entropy Cafe	900, 1st, Floor, Eat Street, 80 Feet Road, Koramangala 6th Block, Bangalore	4.1	Cafes		
Cafe, Beverages, Healthy Food, Juices	Vitamin Palace	21, Hosur Main Road, Chikku Lakshmaiah Layout, Adugodi, Koramangala 7th Block, Bangalore	4.1	Cafes		

Conclusion

- Proposed a content-based restaurant recommendation system using Natural Language Processing (NLP).
- Word Embedding methods: TF-IDF, BERT, Word2Vec, and Countvectorizer to train the recommendation system.
- The supervised proxy problem is implemented by training classification model.
- BERT has outperformed other word embedding models.
- Implement Sentiment Analysis and Named Entity Recognition into the system.
- •A Pre-trained model from DeepNote.com is used for food entity and Part of Speech is used to define the sentiment of review

Limitation

During the sentiment analysis
process, we used the Part of Speech
to define the sentiment of the food
review, but we cannot define the
sentiment perfectly while the
polarity of the review is represented
with multiple sentences.

<u>Challenge</u>

The dataset used in this work is not a labeled dataset as the ground truth to test the training dataset.