Applied Deep Learning for NLP

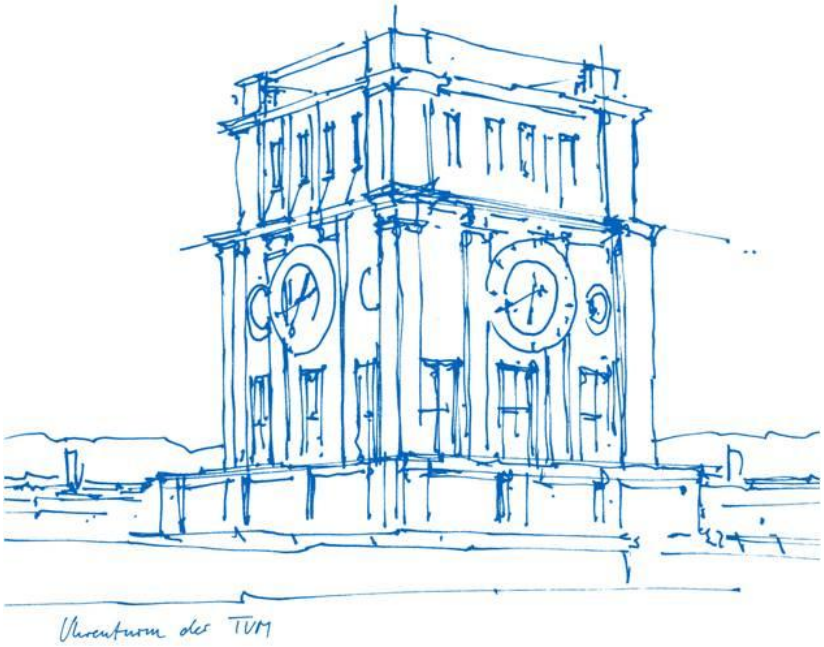
Chon Kit Kuok, Ruiyun Xie

Technische Universität München

Hochschule für Politik

Political Data Science

Munich, 28. January 2021



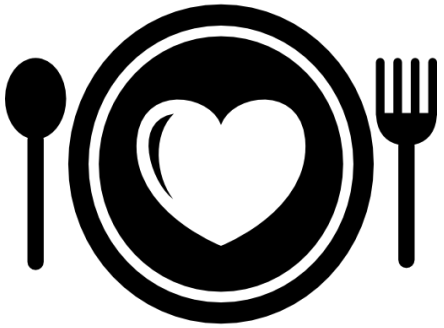
Team

ChonKit Kuok
windygary.kuok@tum.de

Ruiyun Xie
ruiyun.xie@tum.de

Project Idea

- Alexa Skill: Food Match
- Restaurant recommendation based on desired category
- Return review generated with NLP models
- Online smart rating of user review



Dataset

Yelp Open Dataset

Main json files used:

1. business
2. review
3. tip

Categories:

1. Burger
2. Pizza|Italian
3. Vegetarian|Vegan|Salad
4. Chinese|Korean|Asian Fusion|Thai
5. Sushi|Japanese|Ramen
6. French|German|British|Fish & Chips
7. Seafood
8. Mexican|Latin American
9. Steakhouse
10. Others

Models

Text generation model

- Different models for different food categories
- Pretrained GPT2 fitted with new data
- Text generation using combination of Top-K and Top-p sampling
- Five category models
- Used NLPRule for post-processing

Text classification model

- One single DistilBERT model
- smaller and faster version of the BERT
- 100000 positive and 100000 negative reviews for the training
- Rating = score for positive class * 5, range from 0 (lowest) - 5 (highest)

System Architecture

- EC2 for backend server of score prediction model
- Extend server memory by adding swap files in order to load complex model
- Lambda function for skill intents
- S3 (Simple Storage Service) for saving user review

Live Demo

Future Work

- Recommend restaurant based on input location, e.g., Munich
- Add user review and the generated ones to create new dataset

References

- Food Icon
- Yelp Dataset
- NLPRule - Rule-based grammatical error correction library
- Amazon Web Services
- Alexa developer console
- Transformers library