**Computer Science**

**COMP S456F - Software System Development Project**

**Interim Report**

**Project Title: Aspect-Based Sentiment Analysis for Online Restaurant Review in Chinese**

|  |  |
| --- | --- |
| **Student name** | **Student ID** |
| Yeung Ho Yin Tommy | 13024570 |
| Wong Ping Kuen | 13031493 |
| LI Chi Fung | 13031837 |
| Jon Rai | 12749417 |

|  |  |
| --- | --- |
| Group Name | : 2023-Keith-4 |
| Supervisor | : Dr. Keith Lee |
| Submission date | : 27-01-2024 |

**Abstract**

Aspect-Based Sentiment Analysis (ABSA) is one of the popular types of research in deep learning. It focuses on categorizing the aspects and predicting its sentiment polarity from the input. However, the ABSA research usually focused on English or Mandarin (Simplified Chinese). Further, we found that collecting opinions from restaurant review websites is time-consuming. In this report, we create a web application to analyze Chinese (Cantonese) restaurant reviews-based on ABSA research. Specifically, the application aims to save the time collecting opinions from restaurant review websites, obtain more accurate analyzed review results, and also, to handle the lack of ABSA research in Cantonese. To achieve the aims, we created 2 Cantonese restaurant review datasets for the model training, evaluation and applied the ChatGPT 3.5 model to proceed with the data labelling for create the datasets. We created a web application that is used to deploy the trained ABSA model for analyze the OpenRice Chinese restaurant review. Then, we use the test dataset that we created to evaluate our model performance. Lastly, we designed the user satisfaction survey to obtain the opinion from the test user for web application improvement.

**Contents**

[1. Problem Definition 4](#_Toc157267132)

[1.1. Introduction 4](#_Toc157267133)

[1.2. Updated Project Aim 5](#_Toc157267134)

[1.3. Updated Project Objectives 5](#_Toc157267135)

[1.4. Value Propositions 6](#_Toc157267136)

[2. Literature Review (Extended) 6](#_Toc157267137)

[2.1. Related terminologies of Deep Learning 6](#_Toc157267138)

[2.1.1. Accuracy 6](#_Toc157267139)

[2.1.2. F1-Score 6](#_Toc157267140)

[2.1.3. Precision 6](#_Toc157267141)

[2.1.4. Recall 7](#_Toc157267142)

[2.2. Review of Existing or Related Solutions for the Problem 7](#_Toc157267143)

[2.2.1. Aspect-Category Sentiment Analysis (ACSA) 7](#_Toc157267144)

[2.2.2. Aspect-Based Sentiment Analysis (ABSA) 7](#_Toc157267145)

[2.2.3. Existing Models for ACSA and ABSA 8](#_Toc157267146)

[2.2.4. Neural Network used in the existing models 9](#_Toc157267147)

[2.2.5. Models used in existing models 9](#_Toc157267148)

[2.2.6. Existing restaurant review dataset for the Chinese Language 10](#_Toc157267149)

[2.3. Highlight of the proposed solution 11](#_Toc157267150)

[3. Methodology 12](#_Toc157267151)

[3.1. Overview 12](#_Toc157267152)

[3.2. Requirements and Key Technologies 13](#_Toc157267153)

[3.2.1. Function List 13](#_Toc157267154)

[3.2.2. Functional requirements 13](#_Toc157267155)

[3.2.3. Essential key technologies 13](#_Toc157267156)

[3.2.4. Technical Gap 14](#_Toc157267157)

[3.3. System Design 15](#_Toc157267158)

[3.3.1. Use-case diagram 15](#_Toc157267159)

[3.3.2. Architecture or High-Level Design 16](#_Toc157267160)

[3.3.2.1. Component diagram 17](#_Toc157267161)

[3.3.2.2. Data-flow diagram 18](#_Toc157267162)

[3.3.3. Sequence diagram 19](#_Toc157267163)

[3.3.4. User interface design 20](#_Toc157267164)

[3.3.5. Dataset Creation 23](#_Toc157267165)

[3.4. Implementation Issue 25](#_Toc157267166)

[3.5. Evaluation Methods and Design 26](#_Toc157267167)

[3.5.1. Performance Evaluation 26](#_Toc157267168)

[3.5.2. User evaluation 29](#_Toc157267169)

[4. References 30](#_Toc157267170)

[5. Appendixes 32](#_Toc157267171)

[Appendix A. Overview of Project Progress 32](#_Toc157267172)

[Appendix B. Revised Project Plan 32](#_Toc157267173)

[Appendix B. (Group Project) Updated Members’ Roles and Responsibilities 33](#_Toc157267174)

[Appendix C1. (Group Project) Team Member #1's Interim Report 35](#_Toc157267175)

[Appendix C2. (Group Project) Team Member #2's Interim Report 37](#_Toc157267176)

[Appendix C3. (Group Project) Team Member #3's Interim Report 39](#_Toc157267177)

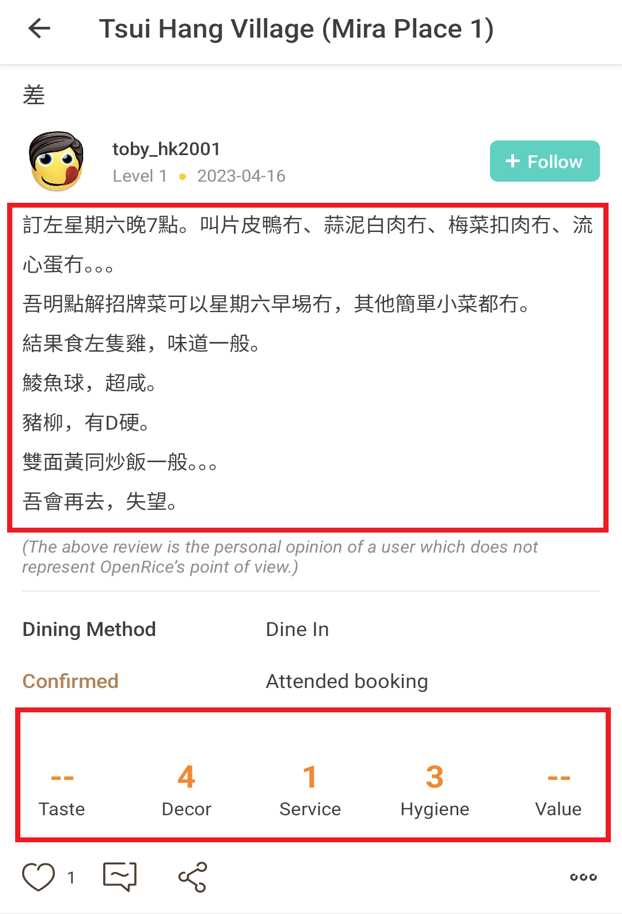
[Appendix C4. (Group Project) Team Member #4's Interim Report 41](#_Toc157267178)

[Appendix D. (Group Project) Updated Meeting Minutes 43](#_Toc157267179)

# Problem Definition

## Introduction

As the internet is becoming more a part of our life, an increasing population are using the internet to collect and generate content, therefore various websites appeared that are close to our lives such as restaurant review websites (E.g. OpenRice and KeeTa, etc.). The restaurant review websites are providing the restaurant information, booking service, review, and rating etc. Many people appreciate sharing their experience in the restaurant in several aspects on these websites about the food quality, service, price etc. Therefore, restaurant customers could read the reviews and rating scores to look for suitable restaurants. Furthermore, restaurant owners could easily collect opinions from the clients for improving their restaurant. However, we found that most of the meaningful comments are always articles which are time-consuming to read.



**Figure 1: A user comment screen captured on OpenRice** (toby\_hk2001, 2023)**, the upper rectangle shows the reviewer’s review about the restaurant, and the lower rectangle shows the rating for each aspect in restaurant by reviewer.**

Moreover, we found that some of the review-mentioned aspects are not reflected in the rating. For example, in Figure 1, the reviewer complained the food quality overall was too bad, but he/she did not provide a rating score for food quality, which only provides the score of Decor (Environment), Service, and Hygiene. Furthermore, Aspect-Based Sentiment Analysis (ABSA) is popular research in natural language processing in Artificial Intelligence (AI), its aims to categorize the aspects from the given input and predicting its polarity in positive, negative, and neutral (Zhang, Zhou, & Wang, 2022). But we found that most of the ABSA research still focuses on English instead of Chinese (Cantonese), which is lack of Chinese (Cantonese) resources. Therefore, a tool for summarizing the polarity of each aspect from the customer’s review text is especially important.

## Updated Project Aim

The project aim of this project is changed to the following:

This project aims to develop an Aspect-Based Sentiment Analysis (ABSA) tool named “Restaurant Review Analyzer for Chinese Language (RRACL)” for restaurant owners to better understand his/her restaurant(s) which parts need to be improved and for restaurant customers to see more accurate rating scores of the restaurant. This tool could also solve the problem of the lack of ABSA research in the Chinese (Cantonese) language.

The reason behind the change is due to that we found the Aspect-Based Sentiment Analysis (ABSA) model can classify the aspects and predict its sentiment polarity more detailed than Aspect-Category Sentiment Analysis (ACSA) model, also we have found that a better ABSA model that directly supported the Chinese language (Only Simplified Chinese). A discussion was held between me and the supervisor on 09-01-2024 and the change was approved.

## Updated Project Objectives

The project objectives of this project is changed to the following:

In our project, we will develop a web application that able to analyze the restaurant review in Chinese (Cantonese) language and provide analyzed results. To achieve these aims, we are using the divide and conquer method to separate the development process into 7 objectives. The details are listed below:

* To collect Chinese review data (raw data) from HK restaurant review websites (E.g., OpenRice) for the preparation of building Chinese datasets.
* To cleanse the collected data (raw data) from the restaurant review website to ensure the dataset's accuracy and correctness (What is Data Cleansing?, n.d.).
* To label the cleansed data in each aspect to building the dataset.
* To adapt the existing models of Aspect-Based Sentiment Analysis (ABSA) for Simplified Chinese to handle the Chinese (Cantonese) dataset.
* To train the existing models of Aspect-Based Sentiment Analysis (ABSA) with Chinese datasets
* To develop a prototype web application for deploying the trained model to handle the restaurant review analysis task and show the analysis results.
* To initiate a user satisfaction survey via Google Forms to collect opinions from test users for web application improvement.

The reason behind the change is due to the same reason with the project aim. A discussion was held between me and the supervisor on 09-01-2024 and the change was approved.

## Value Propositions

For restaurant owners, the tool “Restaurant Review Analyzer for Chinese Language” can help collect customer reviews and generate statistical reports about each aspect of his/her restaurant. The owner can improve their restaurant based on the above reports, so that the restaurant will attract more customers, and the restaurant table turnover rate will also increase (Agilence Staff, 2023). For restaurant review website’ users, they can quickly see the summarized restaurant polarity and the polarity of each comment for each aspect, and these polarities are more accurately compared with the original score (provided by restaurant review websites). From the market perspective, the markets currently lack the existing tools that deployed the ABSA models in Chinese (Cantonese) language, which means our project develops a product that fewer people have done before.

# Literature Review (Extended)

## Related terminologies of Deep Learning

### Accuracy

In our project, the results of our dataset will yield an accuracy score, which is calculated by dividing the number of correct predictions by the total number of predictions across all classes. However, if the number of positive examples is insufficient, the accuracy score may not be a reliable reference. To illustrate, let's consider a specific example where the dataset of positive Chinese comments, such as "good" or "tasty," is limited. In this case, the accuracy score may not be precise. Conversely, if there are enough positive Chinese comments, the accuracy score becomes a reliable and applicable reference. (Czakon, 2023)

Here is the equation for Accuracy:

### F1-Score

F1-score is a commonly used evaluated matric to measure the performance of a classification model, specifically tasks such as text classification and sentiment analysis. It combines **precision** and **recall** into one metric by calculating the harmonic mean between those two. (Czakon, 2023)

Here is the equation for F1-Score:

### Precision

Precision is defined as the number of true positives divided by the sum of true positives and false positives. This means that precision measures the proportion of correctly predicted positive observations out of the total predicted. In simpler terms, precision focuses on the number of predicted positives. (SarielWang, 2020)

Here is the equation for Precision:

### Recall

Recall is defined as the number of true positives divided by the sum of the true positives and false negatives. In simple terms, recall focuses on actual positives number. (SarielWang, 2020)

Here is the equation of Recall:

## Review of Existing or Related Solutions for the Problem

### Aspect-Category Sentiment Analysis (ACSA)

**Aspect-Category Sentiment Analysis (ACSA)** is part of the natural language processing of deep learning, it focuses on categorizing the aspects in coarse-grained from a given text and predicting its sentiment polarity, its predicted result usually is positive, negative, or neutral (Liang, et al., 2021).

### Aspect-Based Sentiment Analysis (ABSA)

**Aspect-Based Sentiment Analysis (ABSA)** is another method for handling the sentiment analysis in natural language processing, it mainly focuses on categorizing the aspects in fine-grained and predicting its sentiment polarity, its predicted result usually is positive, negative, or neutral (Trisna & Jie, 2022).

The major difference between ACSA and ABSA are coarse-grained and fine-grained. The details are listed below:

**Example sentence: “Although the steak is tasty, but the environment is so dirty.”**

According to (Li, Wei, Zhang, Zhang, & Li, 2019), the **coarse-grained** is an **Aspect Category (AC)** task that finds the aspects implicitly appearing from the input texts. The AC task assigns the found entities to the corresponding aspects based on the pre-defined categories on datasets, and those pre-defined categories are the domain terms. For the restaurant, the domain terms could be Taste, Decor, Service, Hygiene, and Value etc.

For the example sentence, it mentioned the entities of **“steak”** and **“environment”**, therefore it will assign the **“steak”** to **Taste** aspectand **“environment”** to **Hygiene** aspect.

According to (Li, Wei, Zhang, Zhang, & Li, 2019), the **fine-grained** is an **Aspect Term (AT)** task that finds the aspects explicitly appearing from the input texts. Different from the AC task, the AT task assigns the aspects by the terms of the found entities, and it is a model-based method that means the aspects do not need to be pre-defined by the datasets.

For the example sentence, it mentioned the entities of **“steak”** and **“environment”**, therefore the **“steak”** and **“environment"** will become the aspects.

### Existing Models for ACSA and ABSA

According to (Liang, et al., 2021), the **Aspect-aware graph convolutional network (AAGCN)** is an aspect-category sentiment analysis model used to extract aspects from text. This model uses contextual sentiment dependencies as the replacement of aspects categorizing in coarse-grained for graph construction. The process of finding contextual sentiment dependencies involves aspect-aware word, aspect-aware weight, and aspect-aware graph(s).

Aspect-aware words: AAGCN uses the distinct aspect word (E.g. price, environment) as the pivot to find the highly related aspect words from the external knowledge.

Aspect-aware weight: AAGCN educes how the aspect-aware word(s) is important to the corresponding aspect, and uses a Beta Distribution to modelling the appearance probabilities of important aspect-aware words to obtain the weights.

Aspect-aware graph(s): AAGCN uses found aspect-aware words and its aspect-aware weight to construct the aspect-aware graph, and this graph uses to learn the contextual sentiment dependencies.

Moreover, AAGCN has provided the non-BERT models and BERT-based models (Liang, et al., 2021).

However, the development of this model is based on the English datasets, which means it does not support the Chinese dataset, and also this model only provides the model training which does not provide a user interface for deploying the trained model.

According to a recent study, (Zhang, Zhou, & Wang, 2022) proposed a new aspect-based sentiment analysis model for deep learning, named **Syntactic and Semantic Enhanced Graph Convolutional Network (SSEGCN).** This model uses syntactic and semantic for graph-base learning task. The model contains several components to achieve the goal.

Contextualized Word Representations, contextualized representations of words are captured by using sentence encoder, it helps to understand the meaning of each word within the sentence context.

Aspect-Aware Attention, semantic correlations that related to different aspect terms or aspect of interest in the sentence are capture, enhance this mechanism combines with self-aware to learn both aspect-related and global semantic information effectively.

Syntactic Mask Matrices are constructs according to the distances between words in the sentence’s syntactic dependency structure to calculate the syntactic mask matrices. Combining

Both adjacency matrices and syntactic mask matrices are to enhance the GCN, allowing the model to fully utilize both syntactic and semantic. It will enhance the understanding and representation of sentences.

Moreover, SSEGCN also has provided the non-BERT models and BERT-based models (Zhang, Zhou, & Wang, 2022).

However, this solution has the same problem as the AAGCN model which is the model only supports the English datasets and it does not provide a user interface for deploying the trained model.

AI Challenger is a platform in China that has hosted artificial intelligence (AI) challenges and competitions since 2017. The challenge covers natural language processing, language translation, weather forecasting, etc. That platform provides a lot of datasets, benchmark models and evaluation metrics to development of AI technologies. Sentiment Analysis is one of the 2018 challenges in AI challenger. **Fine-grained Sentiment Analysis of User Reviews** is the title of ranked 4th team. The project used three layers of transformer to let input data transform to index (integer) for pre-training data, then put that data to convolutional neural networks for training or prediction data. The project can predict simple or transitional Chinese data since that used “opencc” model translates data to simple or transitional Chinese (JohanyCheung, 2018).

Cantonese data also can be predicted, but the project model does not completely predict Cantonese data since the model’ dataset does not have enough data to support Cantonese.

### Neural Network used in the existing models

**Convolutional Neural Networks (CNNs)** are the type of feed-forward network that mainly learns engineering by itself. So, CNNs are a fully connected network, where each neuron in one layer is connected to all neurons in the next layer. CNNs use a combination of three convolutional layers, these are convolutional layers, pooling layers and fully connected layers. CNNs are commonly used in image and video recognition, image segmentation, image classification etc. CNNs can be used in natural language processing, but CNNs are more successful and efficient in image and video recognition (O'Shea & Nash, 2015).

**Graph Convolutional Network (GCN)** is based on Convolutional Neural Network (CNN) for development that operates on graph-structured data. GCN has two generations, there are Spectral Networks and Locally Connected Networks on Graphs (Bruna, Zaremba, Szlam, & LeCun, 2014), and Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering (Defferrard, Bresson, & Vandergheynst, 2016). Spectral Networks and Locally Connected Networks on Graphs is the first generation and is based on the spectrum of its graph-Laplacian. Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering is a second generation and based on bridging the gap between signal processing and spectral graph theory. Those GCN is good at graph analysis, it just little paper is for text analysis with classification. Also, it not completely supporting Chinese language analysis and cannot handling noisy.

### Models used in existing models

According to the survey of (Trisna & Jie, 2022), **Global Vectors for Word Representation (GloVe)** is one of the models that is used for ACSA and ABSA tasks, numerous researchers use GloVe as their word embedding. The GloVe model uses co-occurrence matrix that counts the number of words that appear in context, it factorizes the matrix to obtain the word vectors. GloVe can obtain word embedding from the restaurant's comments then represent the words in a high-dimensional space. Words with similar meanings are closer together, identifying the aspect in the comment and sentiment associated with each aspect.

According to the survey of (Trisna & Jie, 2022), **Word Embedding** is a type of word representing a word with similar meanings to have similar representation. Words are distributed to represent text, one of the keys to impressive performance in deep learning methods in natural language processing problems (NLP). The words represent vector capture semantic and syntactic between words through this method is more efficient and accurate processing in natural language data. Neural network models, such as Word2Vec or GloVe generate word embedding to learn to predict the context of words that appear in a large corpus of text.

According to the survey of (Trisna & Jie, 2022), **Bidirectional Encoder Representations from Transformers (BERT)** is another model designed to understand pretrained deep bidirectional representations by considering both the left and right text. BERT learns to improve its understanding of individual words by relationship between words in a sentence-level context. BERT has achieved state-of-the-art result in natural language processing task, leading to developer to use it on many other pre-trained language models.

According to the article of (aditya\_taparia, 2023), **Bidirectional LSTM(BiLSTM)** is a model or architecture for sentiment analysis to use include Aspect-Based Sentiment Analysis (ABSA) or Aspect-Category Sentiment Analysis (ACSA). BiLSTM is a recurrent neural network technique which is for input in the forward direction and other is for processing in the backward , through this technique to find out the relationship between the word which the word is in the dataset or data, after checking the relationship with the word, the BiLSTM will analyze the word or data and that will output a result or classify the word or data include which sentence is positive or negative.

### Existing restaurant review dataset for the Chinese Language

**Aspect category Sentiment Analysis and rating Prediction (ASAP)** is a large-scale Chinese restaurant review dataset and used for the training dataset of ACSA (Bu, et al., 2021). This dataset collected 46,730 comments real restaurant reviews from the e-commerce platform in China, and each review has assigned 5 coarse-grained aspects for those comments: food, service, price, ambience and miscellaneous (Bu, et al., 2021). However, this dataset has the limitation that is only supported by the Simplified Chinese Language.

## Highlight of the proposed solution

The following table is a comparison of our solution and existing model.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | AAGCN-ACSA | SSEGCN-ABSA | Fine-grained Sentiment Analysis of User Reviews from AI Challenger | Our Solution |
| Find the aspects in coarse-grained / fine-grained | Coarse-grained | Fine-grained | Fine-grained | Fine-grained |
| Neural network | GCN | GCN | CNN | CNN |
| Supported Models | non-BERT models with GloVe / BERT-based models | Non-BERT models / BERT-based models | Transformer+Convolutional | Only support AI Challenger Transformer+Convolutional models |
| Support Chinese Dataset | No | No | Yes (only Simplified Chinese) | Yes (Simplified Chinese and Chinese (Cantonese translate to Simplified Chinese)) |
| Training Model | Yes | Yes | Yes | Yes |
| User Interface for model deployment | No | No | No | Yes |

**Table 1: Solution Comparison**

|  |  |  |  |
| --- | --- | --- | --- |
| Model | Dataset | Accuracy | F1 |
| AAGCN-ACSA | REST15 | 82.79 | 67.43 |
| AAGCN-ACSA-BERT | REST15 | 87.92 | 71.75 |
| SSEGCN-ABSA | REST14 | 84.72 | 77.51 |
| SSEGCN-ABSA-BERT | REST14 | 87.31 | 81.09 |
| Fine-grained Sentiment Analysis of User Reviews from AI Challenger | AI Challenger Sentiment Analysis Training set 2018 | Not mentioned | 72.6 |
| Our Solution | AI Challenger Sentiment Analysis Training set 2018 + our Chinese Dataset (Cantonese) | To be evaluate after completed | To be evaluate after completed |

**Table 2: Solution Performance Comparison**

# Methodology

## Overview

Based on the documents we have researched, analyzing outsourcing and identifying the issues we have solutions to it. By creating a web application, “Restaurant Review Analyzer for Chinese Language” (RRACL) is a web application that provides visualized output like statistic charts, the users can use on any device as long the device is connected to the internet. Users can use RRACL to search for a restaurant's score that has been analyzed by using natural language processing technology, by using the controller to retrieve the review of the restaurant from the open rice platform, then the model will handle the retrieved review by analyzing the comments and generate overall polarities of restaurant’s overall performance and aspect. This web application deployed the model (AI Challenger) that directly supports the Chinese dataset. Furthermore, we have collected and cleansed 1200 Chinese (Cantonese) restaurant reviews from OpenRice and applied the prompt technique to obtain classified and labeled results from the ChatGPT 3.5 model for the dataset creation. To evaluate the performance of our web application, we used the created Chinese restaurant dataset to evaluate our model’s performance. Furthermore, we created a user satisfaction survey to obtain user feedback to improve the user experience of our web application.

## Requirements and Key Technologies

### Function List

|  |  |
| --- | --- |
| **User Types** | **Functions** |
| Restaurant’s Customer | Search restaurant  Display restaurant overall polarity  Display each comment and its aspect polarity |
| Restaurant’s Owner | Search restaurant  Display restaurant overall polarity  Display each comment and its aspect polarity  Display restaurant’s statistic chart  Download restaurant’s statistic report |

**Table 3: Function List**

### Functional requirements

|  |  |
| --- | --- |
| Functions | Description |
| Search restaurant | A function provided for users to search restaurants that they want to see in the web dashboard. |
| Display overall score | Display the summarized aspect score for the restaurant. |
| Statistic calculation | A calculation function that calculates the statistic with restaurant comments is positive, negative, or neutral. |
| Display statistic result | Show the result of statistic calculation. |
| Display statistic in chart | Show the result of statistic calculation with chart. |
| Display comment and its aspect polarity | Display a list of restaurant customer comments and its corresponding aspect polarity. |
| Download statistic report | A download function for users to download a statistic report of currently viewed restaurant. |

**Table 4: Functional requirements**

### Essential key technologies

For our application development, the development language will use Python 3.6.5, which is a web-based application that uses the Python Flask web framework (What is Flask Python, n.d.) to develop the application. For the web user interface creation, the Bootstrap 5 front-end framework (Bootstrap 5 Get Started, n.d.) will be used. Moreover, we will use Chart.js (Chart.js, 2023) to create several charts showing statistical results. The hardware requirement is required to use the Nvidia Graphics Processing Unit (GPU) that contains CUDA cores to speed up the process of AI model training (NVIDIA CUDA in AI Deep Learning, 2022). For the software library, we will use Hugging Face (Lutkevich, 2023) to find NLP resources that benefit our application development. For the model, we will adapt and deploy the existing model Fine-grained Sentiment Analysis of User Reviews from AI Challenger (JohanyCheung, 2018) to handle Chinese (Cantonese) and Simplified Chinese datasets. For the Python modules, we will use TensorFlow (Yegulalp, 2024) as the deep-learning module to build the Aspect-Based Sentiment Analysis (ABSA) models. We also use the Beautiful Soup module (What is Beautiful Soup?, n.d.) to collect review data on the website for dataset creation. For the dataset creation, we will use the ChatGPT 3.5 model combined with the ChatGPT prompt technique (Tamsin, 2023) to help us classify collected data aspects and predict their polarities. Moreover, we will use the OpenCC module (BYVoid, 2023) to translate the Chinese (Cantonese) data to Simplified Chinese data for dataset creation and handle the model's input to ensure the model returns more accurate results. For the model evaluation, we will the scikit-learn module (Metrics and scoring: quantifying the quality of predictions, n.d.) to evaluate the model’s accuracy and F1 score.

### Technical Gap

We found that most of the Aspect-Based Sentiment Analysis (ABSA) models like Fine-grained Sentiment Analysis of User Reviews from AI Challenger (JohanyCheung, 2018) only provide the functionality of models training and models testing, which means it does not provide a user interface for deploying the model. Furthermore, although this model is developed for the dataset of the Chinese Language, it only supports Simplified Chinese, which meaning it cannot directly process Chinese (Cantonese) text. Moreover, although there are Chinese datasets like the Aspect category Sentiment Analysis and rating Prediction (ASAP) (Bu, et al., 2021), but it only provides Simplified Chinese review data.

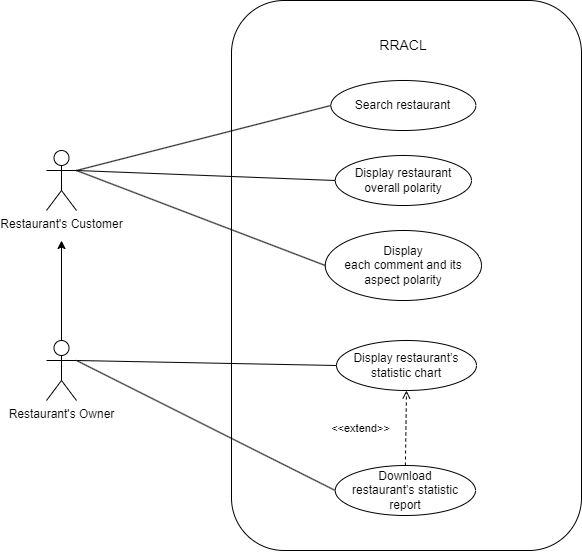
To address these problems, we will make a web application for deploying an ABSA model that can process the Chinese (Cantonese) text and this model is trained by a Chinese (Cantonese) dataset. The application named “Restaurant Review Analyzer for Chinese Language (RRACL)”. The development details are listed below:

* Stage 1: To uses the software library “Hugging Face” to find the NLP resources.
* Stage 2: Manually collect the suitable data from the HK restaurant review website (E.g., OpenRice), cleanse the invalid data, and use ChatGPT 3.5 model with ChatGPT prompt technique to classify the aspects and its polarities. Uses the OpenCC module to translate the review data from Chinese (Cantonese) to Simplified Chinese and add the result (review data and classified results) to the Cantonese dataset (training and evaluation).
* Stage 3: Adapt the code of the AI Challenger model to handle our created Chinese review dataset. (The implementation of AI Challenger model uses the TensorFlow)
* Stage 4: Try to train the models with the created datasets using GPU and evaluate the models’ performance by scikit-learn module with evaluation dataset that we created.
* Stage 5: Use Flask web framework to develop a web application for model deployment and web server deployment.
* Stage 6: Write a script with the Beautiful Soup module to collect the data from the HK restaurant review website (E.g., OpenRice) and use the OpenCC module to convert the collected from to Simplified Chinese as the model’s input.
* Stage 7: Use Bootstrap 5 front-end framework and Chart.js visualization libraries to create user interfaces for showing the model’s results in the board and chart in format. Besides that, it will show a list of the customer comments as well.

## System Design

## Use-case diagram

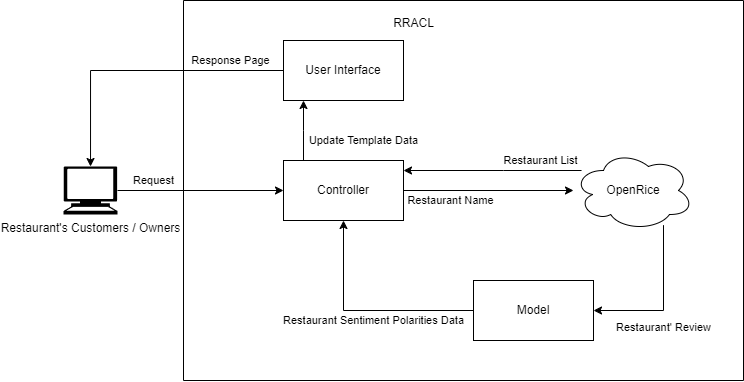
The figure below shows the use case of Restaurant Review Analyzer for Chinese Language (RRACL). In our system, there are 2 target users, which are Restaurant’s Customer and Restaurant’s Owner. For the Restaurant’s Customer, he/she can use this application to search for a target restaurant to overview its overall sentiment polarity (in positive, negative, and neutral), also can view each comment with its aspects’ sentiment polarity. For the Restaurant’s Owner, he/she can having advanced functionalities in this application like show the analyzed results in Charts and download the statistic results report for restaurant improvement.



**Figure 2: Use Case of the system Restaurant Review Analyzer for Chinese Language (RRACL)**

## Architecture or High-Level Design

The figure below shows the Restaurant Review Analyzer for Chinese Language (RRACL) architecture.

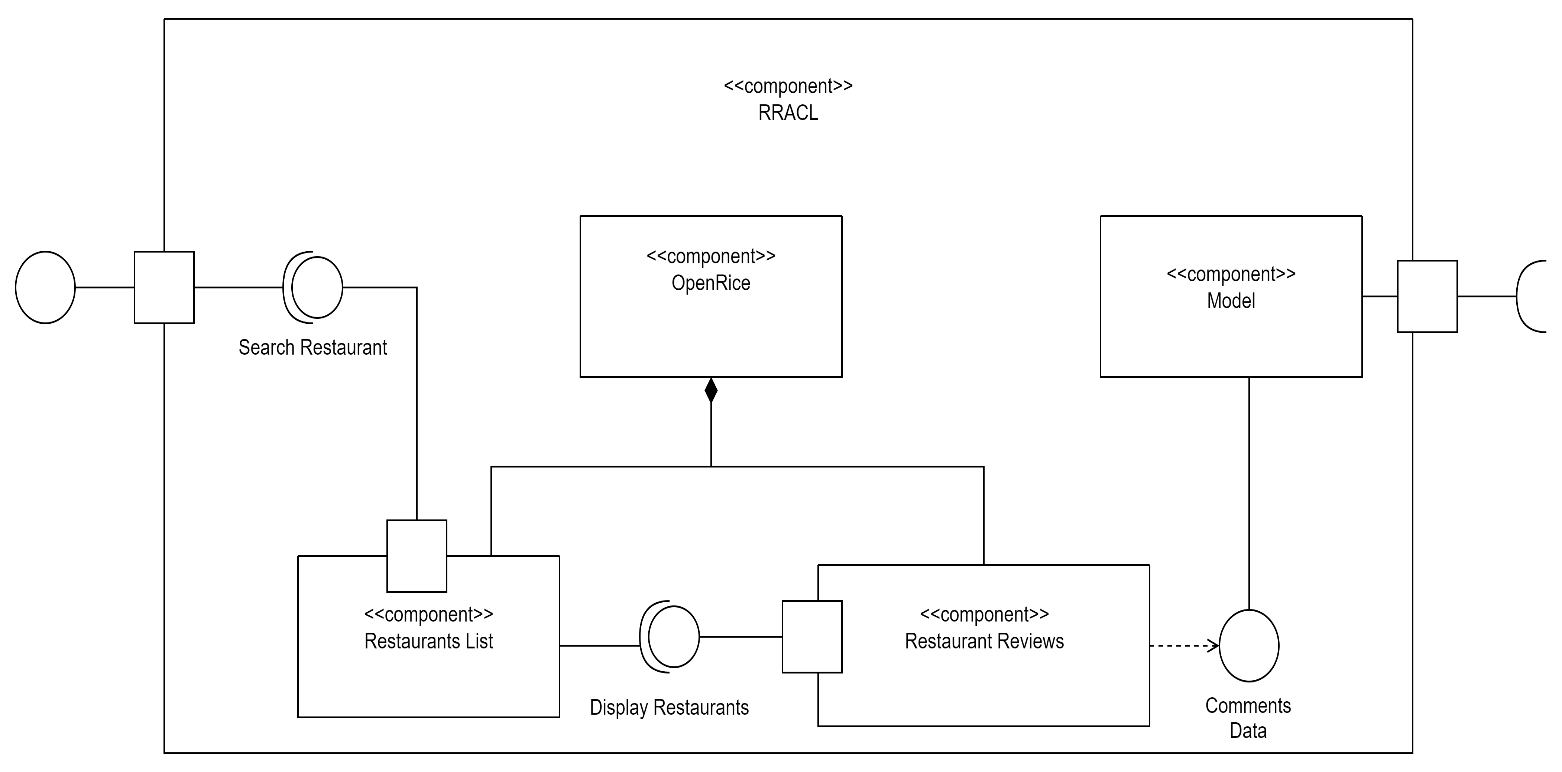


**Figure 3: High-Level System Design of the system Restaurant Review Analyzer for Chinese Language (RRACL)**

* **Controller**: It handles user requests, processes the data from the Model or “Open Rice” website, and updates User Interface and page data.
* **User Interface**: It responds to web pages and displays data to the user.
* **Model**: It is the Aspect-Based Sentiment Analysis (ABSA) model. It handles restaurant reviews from the “Open Rice” website and outputs the restaurant Sentiment Polarities data.
* **OpenRice**: It is the restaurant information and review website.

### Component diagram

The figure below shows the component of Restaurant Review Analyzer for Chinese Language (RRACL).

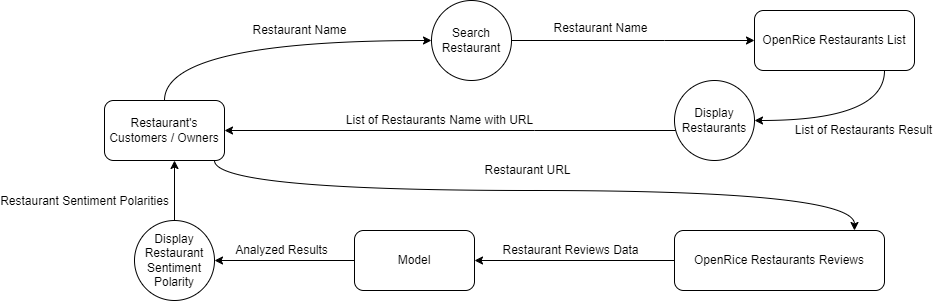


**Figure 4: Component of the system Restaurant Review Analyzer for Chinese Language (RRACL)**

* **RRACL**: It is system’s controller handles user requests, processes the data from the Model or “Open Rice” website, and updates User Interface and page data.
* **Search Restaurant**: It is searching web page. It handles user input the restaurant’s name and request to “OpenRice”’ “Restaurants List”.
* **OpenRice**: It is the restaurant information and review website.
* **Restaurants List**: It is the OpenRice restaurants' list. It will return a list of restaurants that the range be like user input restaurant’s name.
* **Display Restaurants**: It displays search results web page. It handles user select the restaurant’s name and request to “OpenRice”’ “Restaurants Reviews”.
* **Restaurants Reviews**: Is the OpenRice restaurants' list. It will return a list of restaurant reviews that the user selected restaurant.
* **Comments Data**: A list of restaurant review data returned by the OpenRice website, and these data will pass to our model after pre-processing.
* **Model**: It is the Aspect-Based Sentiment Analysis (ABSA) model. It will return analyzed results that the OpenRice got reviews.

### Data-flow diagram

The figure below shows the dataflow of Restaurant Review Analyzer for Chinese Language (RRACL).



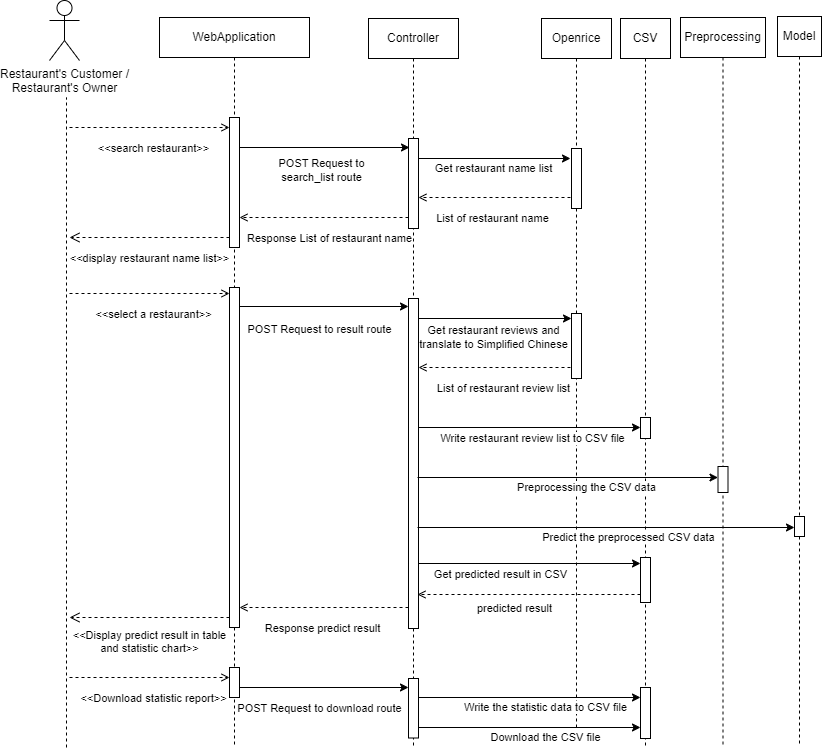
**Figure 5: Data-flow of the system Restaurant Review Analyzer for Chinese Language (RRACL)**

|  |  |
| --- | --- |
| Restaurant's Customers / Owners | Is the system user. He/she will input the restaurant's name. |
| Restaurant Name | It is a name given by the users who want to search for find the restaurant. |
| Search Restaurants | It is a function that displays search web pages. |
| Restaurant Name | It is user input and passing it from “Search Restaurant” to “OpenRice Restaurants List”. |
| Open Rice Restaurant List | Is the OpenRice restaurants' list. It will return a list of restaurants that the range be like user input. |
| List of Restaurants Result | A list of restaurant data returned by the OpenRice website, and this data will pass to Display Restaurants function. |
| Display Restaurants | It is a function that displays search results web page. |
| List of Restaurants Name with URL | It is the range of restaurants searched name by user. Each restaurant's name has their URL. |
| Restaurant's Customers / Owners | Is the system user. He/she will select the restaurant's name. |
| Restaurant URL | It is a name selected by the users who want to find the restaurant. |
| OpenRice Restaurants Reviews | It is the OpenRice restaurants' reviews list. It will return a list of restaurant reviews that the user selected restaurant. |
| Restaurant reviews data | A list of restaurant review data returned by the OpenRice website, and these data will pass to our model after pre-processing. |
| Model | It is the Aspect-Based Sentiment Analysis (ABSA) model. It will return analyzed results that the OpenRice got reviews. |
| Analyzed Results | The results given by the model are based on the restaurant review data and passed to the display restaurant sentiment polarity function. |
| Display Restaurant Sentiment Polarity | It is a function that displays the model analyzed results web page. |
| Restaurant Sentiment Polarities | The polarities given by analyzed results are each aspect have their sentiment polarity. |

**Tabel 5: Description of data flow diagram**

## Sequence diagram

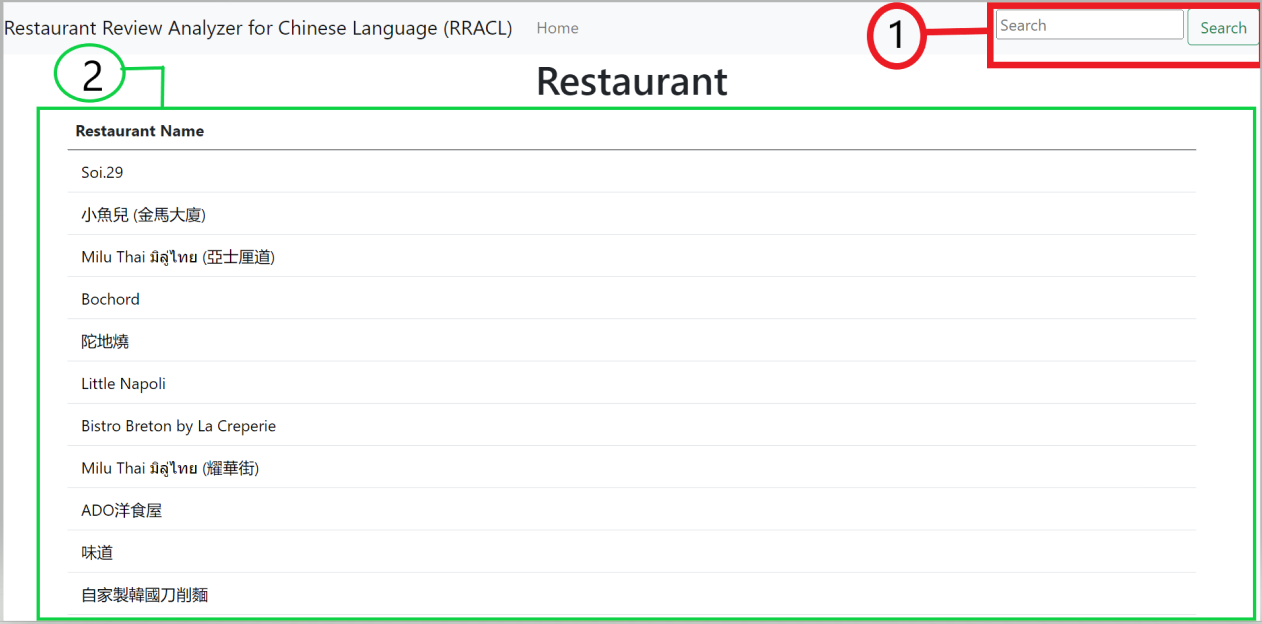
The figure below shows the sequence diagram of Restaurant Review Analyzer for Chinese Language (RRACL).



**Figure 6: Sequence diagram of the system Restaurant Review Analyzer for Chinese Language (RRACL)**

## User interface design

During the website interface design process, we placed significant emphasis on optimizing usability and task suitability. Our primary objective was to ensure that users can easily comprehend and navigate the website, which was achieved through meticulous attention to detail during the graphical user interface (GUI) prototype design phase. We have applied Bootstrap 5, JavaScript and Flask technology to add dynamic user interfaces features. We used JavaScript to create an interactive animation, when user clicked one of the restaurant names to show the detail a loading animation will be displayed Figure 11. The website includes **the Main page,** **Search page** and **result page.**

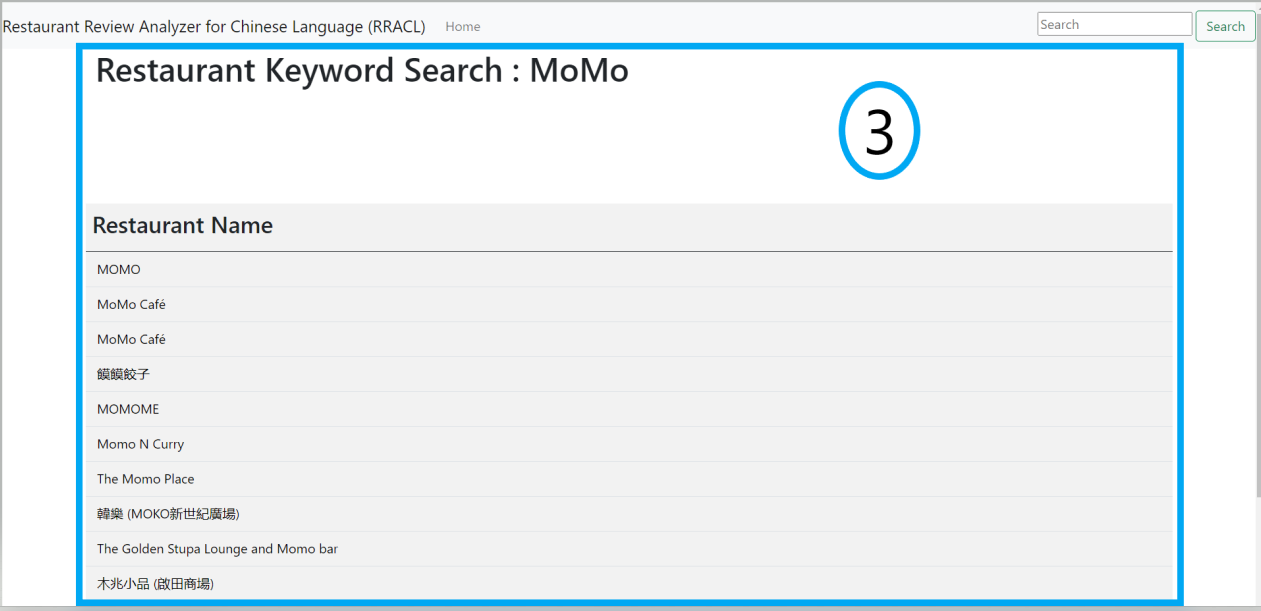


**Figure 7: Main Page will show top 30 restaurant of Open rice.**

The website's main page will list Open Rice's top 30 restaurants Figure 7, along with a search bar Figure 7(1) conveniently located on the navigation bar.

1.) Search Bar for user to search for restaurant name.

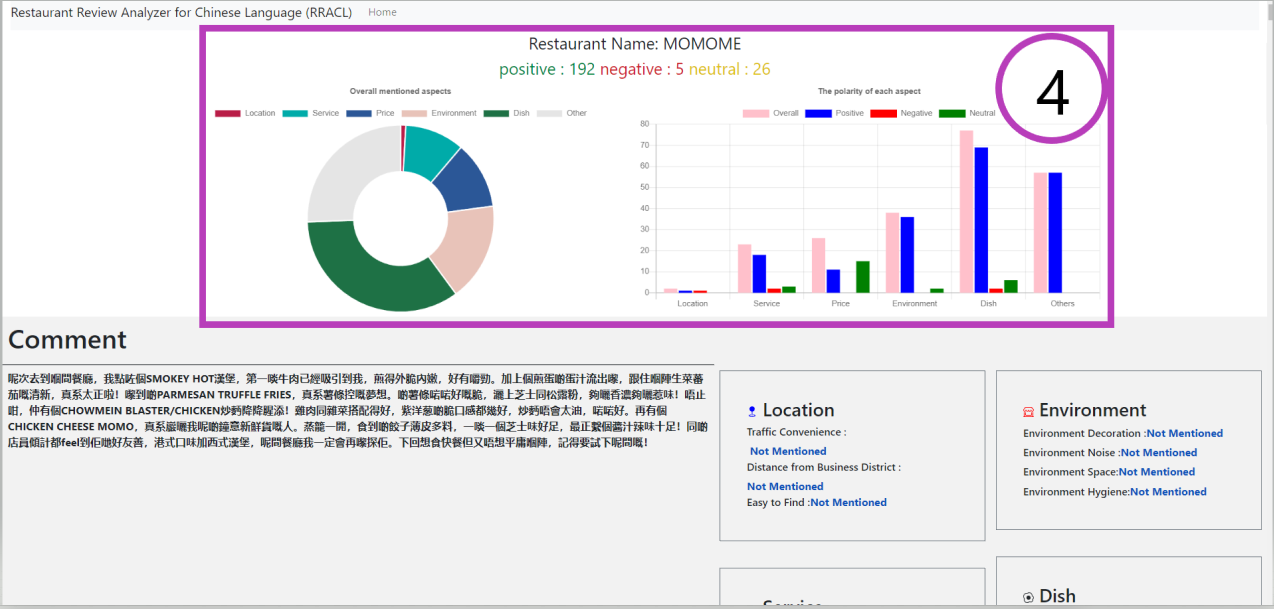
2.) Shows top 30 restaurants from open rice.



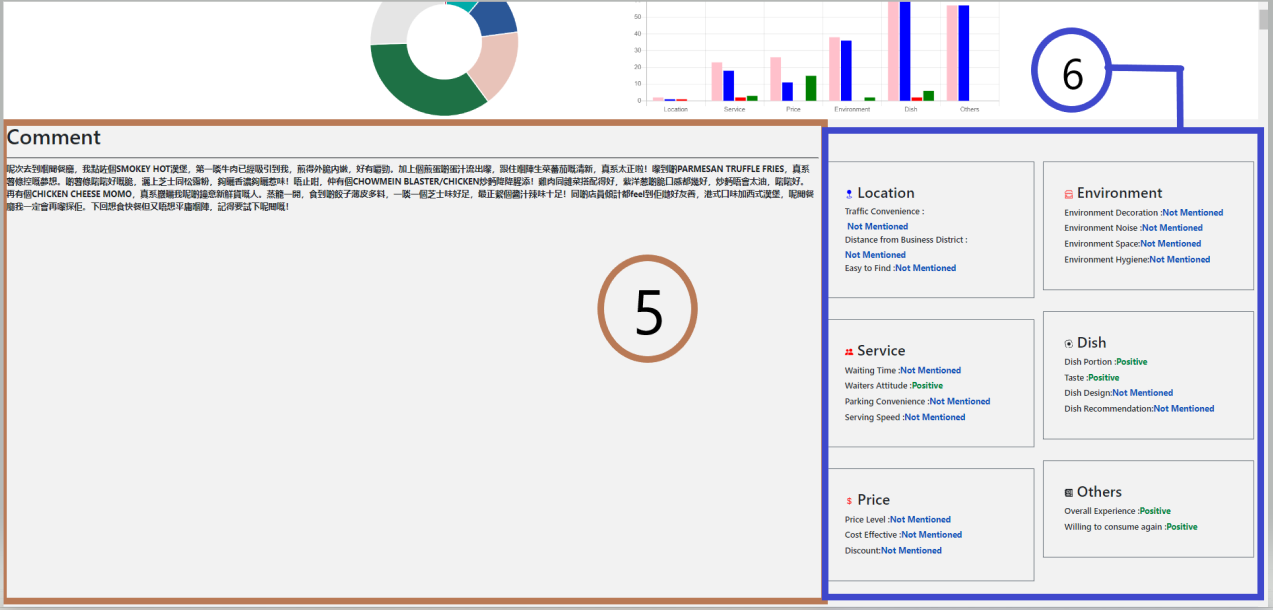
**Figure 8: Search Page displaying result in restaurant detail.**

The Search Page Figure 8 will present the names of restaurants that have been queried by the user through the search bar. The keyword that is used to search will be presented in “Restaurant Keyword Search:” Figure 8(3).

3.) list of restaurants name that was searched.



**Figure 9: Result Page showing restaurant analyze result in graphs.**

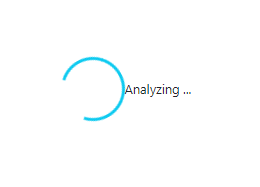
**Figure 10: Result Page showing restaurant's comments and analyze result in Dashboard.**

On the results page Figure 9-4, the user will be able to see analysis result of the restaurant in two graphs, bar chart and doughnut chart. Both graphs Figure 9(4) will show aspects resulting in several positive, negative and neutral. And more detail will be displayed on the comments section Figure 10(5) with analyze result Figure 10(6) according to the comment.

4.) restaurant's result after analyzing the comments with model.

5.) restaurant's comments retrieved from the open rice.

6.) Dashboard/results after analyzing the comments.



**Figure 11: This is the loading animation when user have wait for restaurant analyze results.**

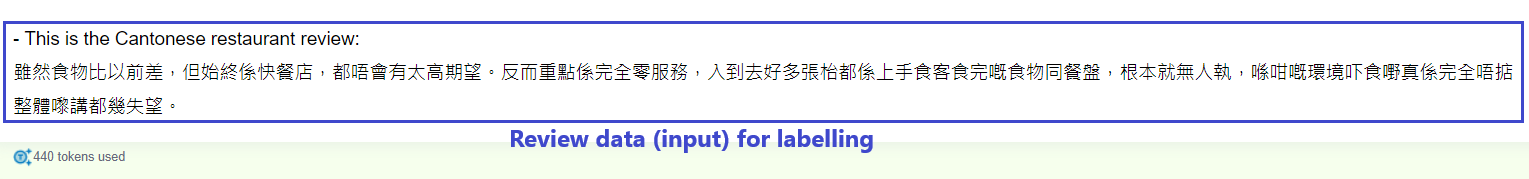
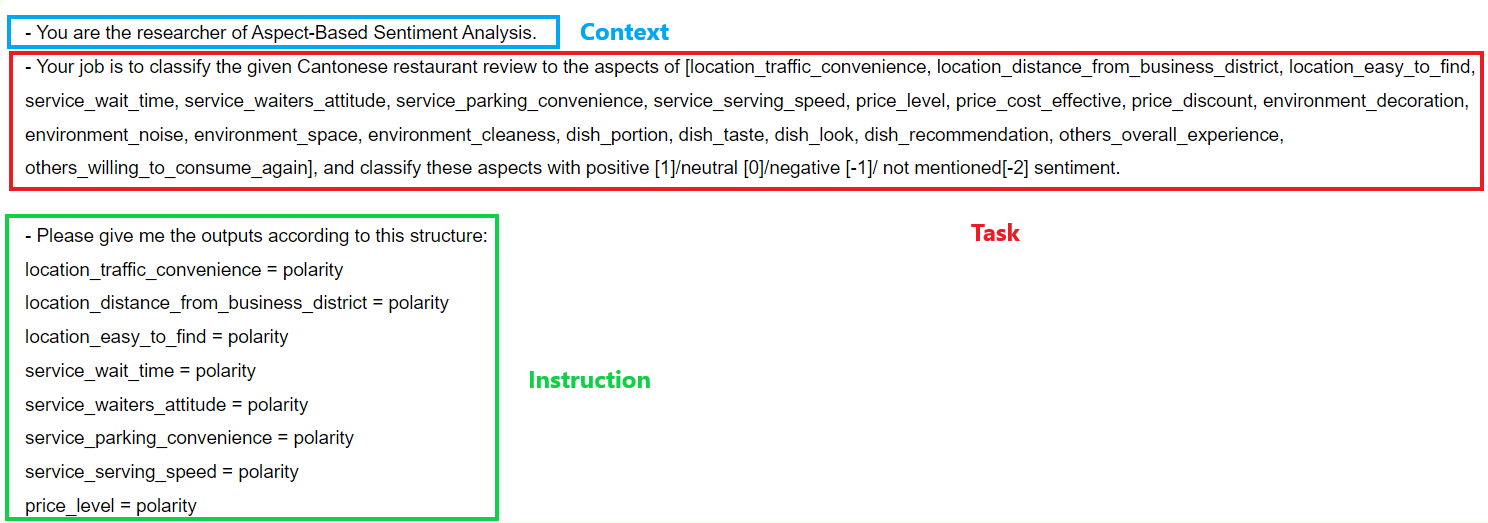
In Figure 11, the loading animation will be presented when the user clicks on one of the restaurants. When the animation/analysis is finished, it will show restaurant details Figure 9 and Figure 10.

## Dataset Creation

We conduct the 2 datasets creation for model training and evaluation. To achieve this, we are separated into 3 major steps which are data collection and cleansing, data labelling, and dataset creation.

* **Data collection and cleansing**: we manually collected 1200 suitable review data (raw data) from OpenRice, cleansed the invalid data like emoji and special characters.
* **Data Labeling**: we use the ChatGPT 3.5 model to classify the review data’s aspects and obtain its sentiment polarity such as positive, negative, neutral, and not mentioned.
* **Dataset creation**: we use the OpenCC module to translate the review data (sentences) to Simplified Chinese and fill the data and the classification result to the CSV file.

The data labelling is the ChatGPT 3.5 model that is provided by HKMU ChatGPT Web Portal and combined with the ChatGPT prompt technique to obtain the labelled data.

**Figure 12: This figure shows the prompt for obtaining the labelled result from ChatGPT 3.5 model.**

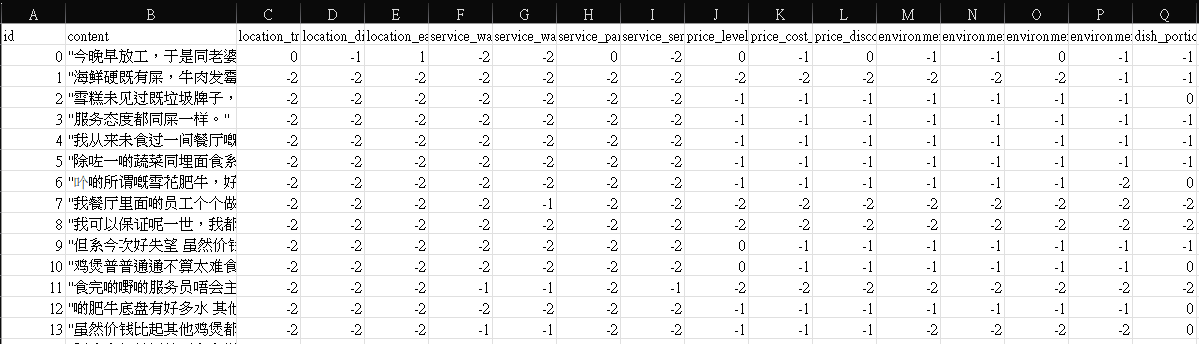
In figure 12, we followed the ChatGPT prompt formula to create our own prompt to obtaining the aspects’ polarities from the given Cantonese restaurant review, the prompt formula usually contained the Context, Task, Instruction, Clarify, and Refine, but only contained the Context, Task, and Instruction in our prompt:

* **Context**: Given the scenario to ChatGPT model for obtaining a specific result, in this case we have given the scenario that is the ABSA researcher.
* **Task**: Given the clear task to ChatGPT to focused on the specified task. In this case, we assigned the task to classify the given Cantonese restaurant review to the provided aspects and classify those aspects’ polarity as positive, negative, neutral, and not mentioned.
* **Instruction**: Given the specific requirements after the provided task, and that requirement ensure the output will follow the instructions, in this case, the point form will be the format of the output.



**Figure 13: It shown result of the labelled data**

Based on the prompt in Figure 12, we can obtain the aspects and its polarities (labelled data) from the Cantonese restaurant review, and the classification result of the ChatGPT 3.5 model shown in Figure 13.



**Figure 14: This figure is screen captured on the CSV dataset that we created; it is the partial data in the dataset.**

After that, we use the OpenCC module to translate the restaurant review to Simplified Chinese, copy the translated data and the result of Figure 13 to the dataset in the shown on Figure 14.

## Implementation Issue

For our prototype, we deployed based on the high-level programming language Python 3.6.5. Then, we use the Beatifulsoup 4.12.2 and lxml 5.1.0 to create a script for scraping the OpenRice restaurant and its review data. For the Aspect-Based Sentiment Analysis (ABSA) model, existing Model of Transformer with convolutional neural network of Fine-grained Sentiment Analysis of User Reviews from AI Challenger with the tensorflow-gpu 1.10.0 module to build the ABSA models. Besides, the model training requires NVIDIA GPU to speed up the training process. For the model evaluation, we use the scikit-learn 0.24.2 module with our created dataset to evaluate the model’s performance. Since the model is trained by the Simplified Chinese dataset which only accepts the Simplified Chinese data, for address this problem, we use opencc 1.1.1 module to translate the scraped review data to Simplified Chinese to ensure the model’s return more accurate prediction results. Furthermore, we use the pandas 0.23.4 module to write the scraped data into CSV file as the model input and use it to read the model’s predicted output in CSV file. To implement web application Flask 2.0.3 and user interface, web framework to develop the web-based application, its user interface and statistic chart based on Bootstrap 5.15.3 front-end framework and Chart.js 2.9.4 visualization libraries to create.

## Evaluation Methods and Design

## Performance Evaluation

We conduct the performance evaluation of our trained Aspect-Based Sentiment Analysis (ABSA) model with our created Chinese (Cantonese) test dataset. This dataset labeled the aspects and their polarities as shown in statistic Table 6. In our ABSA model, we use the AI Challenger Sentiment Analysis Training set 2018 dataset to train our model in the current stage, which is not integrated with our created training dataset. The data details of the training dataset as shown in statistic Table 7.

To conduct our performance evaluation, we use the Python scikit-learn module with our test dataset to evaluate the model. Python scikit-learn has metrics for calculating the model accuracy and F1 score. The accuracy score for each label is produced by setting the predicted data list and correct data list for each label into "accuracy\_score" in "sklearn.metrics" field. The predicted data list and correct data list for each label are entered into "f1\_score" in "sklearn.metrics", and the average of the metrics for each label is set to macro, producing the F1 score for each label.

|  |  |
| --- | --- |
| **Figure 15: The evaluation result of model’s accuracy** | **Figure 16: The evaluation result of model’s F1 score** |

As the result shown in Figure 15 and Figure 16, the accuracy and F1 score of the models are respectively 49.75% and 26.12%. The poor performance of the model is due to several factors. The training dataset lacks Cantonese reviews, although there are some Cantonese reviews that exist in this dataset, but most data in this dataset are Mandarin review data. Furthermore, some of the aspects’ polarity data like **service\_parking\_convenience** lack of positive and negative, and lack of data will significantly affect the model result of evaluation. For improving the model performance, we will append 1500 and 300 labelled Chinese (Cantonese) review data to AI Challenger dataset and the testing dataset to enhance the model’s performance in Cantonese restaurant review.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Our Chinese (Cantonese) test dataset | | | |
| Aspects | Not mentioned | Negative | Neutral | Positive |
| location\_traffic\_convenience | 167 | 7 | 16 | 10 |
| location\_distance\_from\_business\_district | 186 | 4 | 6 | 4 |
| location\_easy\_to\_find | 134 | 8 | 3 | 55 |
| service\_wait\_time | 125 | 32 | 22 | 21 |
| service\_waiters\_attitude | 83 | 42 | 16 | 59 |
| service\_parking\_convenience | 181 | 1 | 15 | 3 |
| service\_serving\_speed | 121 | 17 | 11 | 51 |
| price\_level | 59 | 35 | 61 | 45 |
| price\_cost\_effective | 42 | 37 | 20 | 101 |
| price\_discount | 99 | 18 | 62 | 21 |
| environment\_decoration | 65 | 29 | 65 | 41 |
| environment\_noise | 77 | 35 | 84 | 4 |
| environment\_space | 65 | 26 | 54 | 55 |
| environment\_cleaness | 64 | 26 | 49 | 61 |
| dish\_portion | 20 | 32 | 55 | 93 |
| dish\_taste | 23 | 57 | 23 | 97 |
| dish\_look | 51 | 30 | 72 | 47 |
| dish\_recommendation | 38 | 37 | 26 | 99 |
| others\_overall\_experience | 6 | 84 | 16 | 94 |
| others\_willing\_to\_consume\_again | 14 | 73 | 11 | 102 |

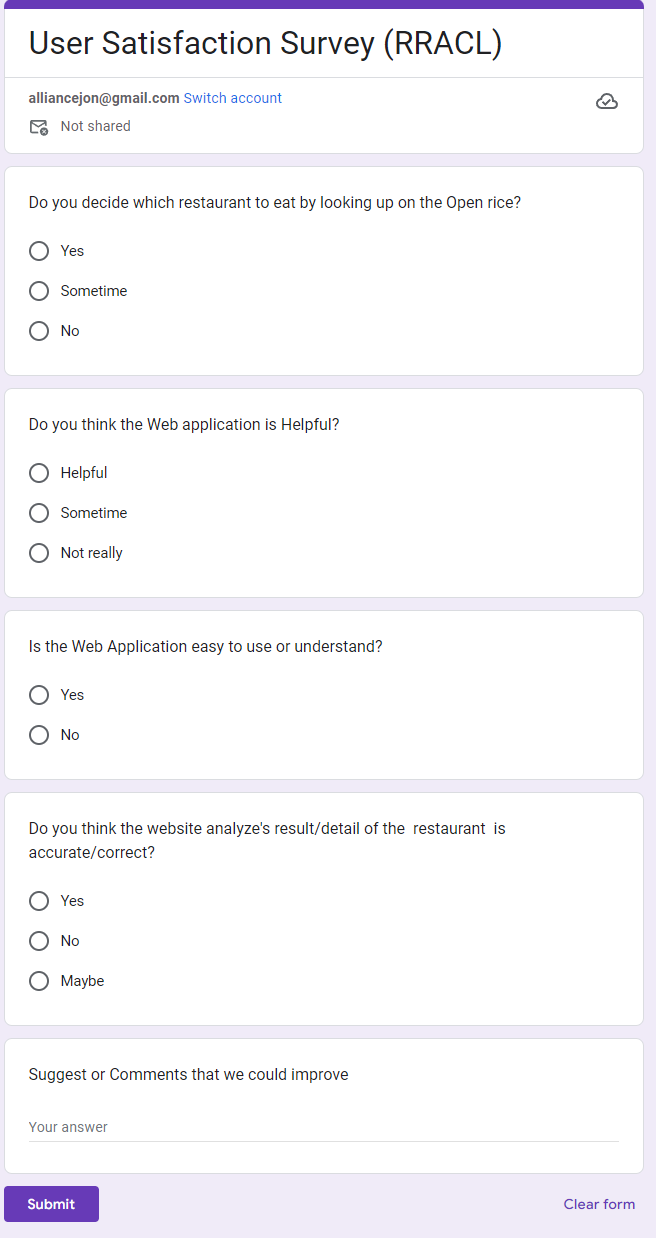
**Table 6: Statistics for the evaluation dataset**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | AI Challenger Sentiment Analysis Training set 2018 (Train) | | | |
| Aspects | Not mentioned | Negative | Neutral | Positive |
| location\_traffic\_convenience | 81382 | 1318 | 1046 | 21254 |
| location\_distance\_from\_business\_district | 83680 | 586 | 533 | 20201 |
| location\_easy\_to\_find | 80605 | 3976 | 2472 | 17947 |
| service\_wait\_time | 92763 | 3034 | 4382 | 4821 |
| service\_waiters\_attitude | 42410 | 8684 | 12534 | 41372 |
| service\_parking\_convenience | 98276 | 1323 | 1456 | 3945 |
| service\_serving\_speed | 88700 | 5487 | 2379 | 8434 |
| price\_level | 52820 | 12375 | 24249 | 15556 |
| price\_cost\_effective | 80242 | 3011 | 3072 | 18675 |
| price\_discount | 64243 | 1716 | 18255 | 20786 |
| environment\_decoration | 53916 | 2139 | 9492 | 39453 |
| environment\_noise | 73445 | 3077 | 4843 | 23635 |
| environment\_space | 65398 | 5706 | 9262 | 24634 |
| environment\_cleaness | 66598 | 4513 | 4703 | 29186 |
| dish\_portion | 56917 | 10018 | 9506 | 28559 |
| dish\_taste | 5070 | 4363 | 40200 | 55367 |
| dish\_look | 75975 | 3178 | 4675 | 21172 |
| dish\_recommendation | 84767 | 2275 | 1988 | 15970 |
| others\_overall\_experience | 2110 | 9384 | 23436 | 70070 |
| others\_willing\_to\_consume\_again | 65600 | 4159 | 2913 | 32328 |

**Table 7: Statistics for the training dataset**

## User evaluation

To evaluate the performance of our prototype web application, we decided to use a survey method, by designing a survey to let the test users fill out the survey on google. The survey will be used to get feedback from users who have tested our system. By using a survey, we can understand which part of our system needs to improve. Figure 17 is our survey design.



**Figure 17: This is our survey to get feedback from users by using google.**

# References

(n.d.).

aditya\_taparia. (2023, June 8). *Bidirectional LSTM in NLP*. (GeeksforGeeks) Retrieved October 10, 2023, from GeeksforGeeks: https://www.geeksforgeeks.org/bidirectional-lstm-in-nlp/

*Bootstrap 5 Get Started*. (n.d.). (W3Schools) Retrieved November 3, 2023, from w3schools.com: https://getbootstrap.com/docs/5.0/getting-started/introduction/

Bruna, J., Zaremba, W., Szlam, A., & LeCun, Y. (2014). Spectral Networks and Locally Connected Networks on Graphs. *arXiv:1312.6203*. doi:https://doi.org/10.48550/arXiv.1312.6203

Bu, J., Ren, L., Zheng, S., Yang, Y., Wang, J., Zhang, F., & Wu, W. (2021). ASAP: A Chinese Review Dataset Towards Aspect Category Sentiment Analysis and Rating Prediction. *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2069–2079. doi:http://dx.doi.org/10.18653/v1/2021.naacl-main.167

*Chart.js*. (2023, April 12). Retrieved from https://www.chartjs.org/docs/latest/

Czakon, J. (2023, September 5). *F1 Score vs ROC AUC vs Accuracy vs PR AUC: Which Evaluation Metric Should You Choose?* (Neptune Labs) Retrieved November 1, 2023, from neptune.ai: https://neptune.ai/blog/f1-score-accuracy-roc-auc-pr-auc

Defferrard, M., Bresson, X., & Vandergheynst, P. (2016). Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering. *30th Conference on Neural Information Processing Systems (NIPS 2016), 29*, 3844-3852. Retrieved from https://proceedings.neurips.cc/paper\_files/paper/2016/file/04df4d434d481c5bb723be1b6df1ee65-Paper.pdf

JohanyCheung. (2018, November 29). *AI Challenger 细粒度用户评论情感分析*. Retrieved from GitHub: https://github.com/JohanyCheung/fsauor/tree/master

Li, Z., Wei, Y., Zhang, Y., Zhang, X., & Li, X. (2019). Exploiting Coarse-to-Fine Task Transfer for Aspect-Level Sentiment Classification. *Proceedings of the AAAI Conference on Artificial Intelligence, 33*(01), 4253-4260. doi:https://doi.org/10.1609/aaai.v33i01.33014253

Liang, B., Su, H., Yin, R., Gui, L., Yang, M., Zhao, Q., . . . Xu, R. (2021). Beta Distribution Guided Aspect-aware Graph for Aspect Category Sentiment Analysis with Affective Knowledge. *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, 208–218.

Lutkevich, B. (2023, September). *What is Hugging Face?* (TechTarget) Retrieved November 3, 2023, from Whatls.com: https://www.techtarget.com/whatis/definition/Hugging-Face

*Metrics and scoring: quantifying the quality of predictions*. (n.d.). Retrieved from scikit learn: https://scikit-learn.org/stable/modules/model\_evaluation.html

*NVIDIA CUDA in AI Deep Learning*. (2022, October 6). (Exxact Corporation) Retrieved November 3, 2023, from EXXACT: https://www.exxactcorp.com/blog/Deep-Learning/NVIDIA-CUDA-in-AI-Deep-Learning

O'Shea, K., & Nash, R. (2015). An Introduction to Convolutional Neural Networks. ArXiv. /abs/1511.08458.

SarielWang. (2020, May 15). *機器學習模型評估指標-confusion matrix, precision, recall, and ,F1-score*. Retrieved November 1, 2023, from Medium: https://medium.com/@s716419/%E6%A9%9F%E5%99%A8%E5%AD%B8%E7%BF%92%E6%A8%A1%E5%9E%8B%E8%A9%95%E4%BC%B0%E6%8C%87%E6%A8%99-confusion-matrix-precision-and-recall-e9d64ff14d81

Staff, A. (2023, May 18). *How To Improve Your Restaurant’s Table Turnover: The Ultimate Guide*. (Agilence, Inc) Retrieved October 3, 2023, from Agilence: https://blog.agilenceinc.com/table-turnover

toby\_hk2001. (2023, April 16). *Tsui Hang Village (Mira Place 1) - Review*. (OpenRice) Retrieved October 10, 2023, from https://www.openrice.com/en/hongkong/review/%E5%B7%AE-e4787429

Trisna, K. W., & Jie, H. J. (2022). Deep Learning Approach for Aspect-Based Sentiment Classification: A Comparative Review. *Applied Artificial Intelligence, 36*(1). doi:10.1080/08839514.2021.2014186

*What is Beautiful Soup?* (n.d.). (Educative, Inc) Retrieved November 3, 2023, from educative: https://www.educative.io/answers/what-is-beautiful-soup

*What is Data Cleansing?* (n.d.). (Amazon Web Services, Inc) Retrieved November 1, 2023, from aws: https://aws.amazon.com/what-is/data-cleansing/

*What is Flask Python*. (n.d.). (pythonbasics.org) Retrieved November 3, 2023, from Python Tutorial: https://pythonbasics.org/what-is-flask-python/

*What is PyTorch?* (n.d.). (NVIDIA Corporation) Retrieved November 3, 2023, from NVIDIA: https://www.nvidia.com/en-us/glossary/data-science/pytorch/

Zhang, Z., Zhou, Z., & Wang, Y. (2022). SSEGCN: Syntactic and Semantic Enhanced Graph Convolutional Network for Aspect-based Sentiment Analysis. *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 4916–4925. doi:http://dx.doi.org/10.18653/v1/2022.naacl-main.362

# Appendixes

## Appendix A. Overview of Project Progress

Tasks Completed and Tasks Ongoing:

The following lists the tasks completed:

* Trained the AI Challenger model with Chinese datasets provided by AI Challenger (Objective 5).
* Developed a prototype web application for deploying the trained model and able to obtain the analyze result from the model based on selected restaurant’s reviews (Objective 6).

The following lists the ongoing tasks:

* Creating 2 Chinese (Cantonese) datasets for evaluation and training model, these datasets have contained 1200 data that already have cleansed and labeled the data. But the datasets are based on the AAGCN model to create, that means it cannot directly apply to the AI Challenger model, it needs some modification. It is about 60% completed.
* Developing the functionality of download statistic report. It is about 10% completed.
* Implementing the user satisfaction survey to collect opinions about the prototype web application. It is about 10% completed.

## Appendix B. Revised Project Plan

|  |  |
| --- | --- |
| Date | Description |
| Week 0 – 19 | Research the existing model and test its code |
| Week 10 – 25 | Prepare the Chinese dataset |
| Week 19 – 25 | Train and evaluate the models |
| Week 19 – 26 | Build web application for model deployment |
| Week 27 – 28 | Collect opinion with survey for web application improvement |
| Week 28 – 33 | Finish the Final product and Final report |

## Appendix B. (Group Project) Updated Members’ Roles and Responsibilities

**Roles**

|  |  |  |
| --- | --- | --- |
| Roles | Member(s) | Remarks |
| Team Coordinator | Yeung Ho Yin Tommy,  Rai Jon |  |
| Secretary | Yeung Ho Yin Tommy | Meeting agenda and minutes |
| Data Analyst | Yeung Ho Yin Tommy,  LI CHI FUNG |  |
| Programmer | Yeung Ho Yin Tommy,  Wong Ping Kuen,  LI Chi Fung,  Jon Rai |  |
| System Analyst and Designer | Yeung Ho Yin Tommy,  Wong Ping Kuen,  LI Chi Fung,  Jon Rai |  |
| Tester and Evaluator | Yeung Ho Yin Tommy,  Wong Ping Kuen,  LI Chi Fung,  Jon Rai |  |
| Artificial Intelligent Expert | Yeung Ho Yin Tommy,  Wong Ping Kuen, | Handle AI model with TensorFlow |
| User Interface Expert | LI Chi Fung,  Jon Rai | Handle Flask web framework |

**Responsibilities**

|  |  |  |
| --- | --- | --- |
| Tasks | Responsible Member(s) | Target Date |
| Technology Test: AI model | Yeung Ho Yin Tommy,  Wong Ping Kuen,  LI Chi Fung,  Jon Rai | 10 Jan 2024 |
| Datasets creation: data collection and data cleansing (Stage 1) | Yeung Ho Yin Tommy,  Wong Ping Kuen,  LI Chi Fung,  Rai Jon | 29 Feb 2024 |
| Datasets creation: Label the aspects for each review (Stage 2) | Yeung Ho Yin Tommy,  Wong Ping Kuen,  LI Chi Fung,  Rai Jon | 29 Feb 2024 |
| AI: Model Adaption (Stage 1) | Yeung Ho Yin Tommy,  Wong Ping Kuen | 15 Mar 2024 |
| AI: Model Training (Stage 2) | Yeung Ho Yin Tommy,  Wong Ping Kuen | 15 Mar 2024 |
| AI: Models Evaluation (Stage 3) | Yeung Ho Yin Tommy,  Wong Ping Kuen | 15 Mar 2024 |
| Web Application: UI design (Stage 1) | Li Chi Fung,  Rai Jon,  Yeung Ho Yin Tommy | 30 Mar 2024 |
| Web Application: UI implementation (Stage 2) | Li Chi Fung  Rai Jon | 30 Mar 2024 |
| Web Application: Web Server Setup (Stage 3) | Yeung Ho Yin Tommy  Wong Ping Kuen | 20 Jan 2024 |
| Web Application: User satisfaction survey (Stage 4) | Rai Jon | 30 Mar 2024 |
| Web Application: Data Analysis (Stage 5) | Wong Ping Kuen,  Yeung Ho Yin Tommy | 1 Apr 2024 |
| Survey design | LI Chi Fung  Rai Jon | 20 Jan 2024 |

## Appendix C1. (Group Project) Team Member #1's Interim Report

Yeung Ho Yin Tommy (13024570)

Project Team Name: 2023-Keith-4

Project Title: Aspect-Based Sentiment Analysis for Online Restaurant Review in Chinese

Supervisor: Dr. Keith Lee

I, (Yeung Ho Yin Tommy, 13024570), certify that the description and information included in this team member's report is true to the best of my knowledge. During October 2023 to November 2023, I trained and evaluated some models such as “AAGCN” and “SSEGCN”. In December 2023, I finished some of the server-sides and dataset creation tasks, such as scraping the review data from OpenRice as the model's input, and the dataset creation with 300 restaurant review data. Furthermore, I studied the ChatGPT prompt technique and wrote a suitable prompt to obtain the classification result for data labelling. However, during the model deployment, I found that the models of “AAGCN” and “SSEGCN” couldn’t deploy to the web application because the data preprocessing script caused the dimension errors, I tried to find a solution to fix this problem till to the earlier of January 2024, but it still did not work. Moreover, I also tried to write a script to connect Poe’s ChatGPT 3.5 model to handle the ACSA problem, but I found is problem in Poe which is Chat limitation. The free account of Poe is only able to send 100 messages to the ChatGPT 3.5 model, so it is not suitable for our project. In January 2024, we found a new ABSA model that directly support Chinese. I trained the model and deployed it to our web application, also I designed the user interface like the result page. Then, I made the statistic function to counting the model's output, and created the statistic charts in the web user interface to display the output of the statistical function. As the created dataset format is not suitable for AI Challenger model, therefore I adapted the previously created dataset to evaluate the AI Challenger model.

Yeung Ho Yin Tommy, 21 January 2024

Task

|  |  |  |  |
| --- | --- | --- | --- |
| Tasks | Responsible Member(s) | Target Date | Completion |
| Technology Test: AI model | Yeung Ho Yin Tommy,  Wong Ping Kuen,  LI Chi Fung,  Jon Rai | 10 Jan 2024 | 100% |
| Datasets creation: data collection and data cleansing (Stage 1) | Yeung Ho Yin Tommy,  Wong Ping Kuen,  LI Chi Fung,  Rai Jon | 29 Feb 2024 | 60% |
| Datasets creation: Label the aspects for each review (Stage 2) | Yeung Ho Yin Tommy,  Wong Ping Kuen,  LI Chi Fung,  Rai Jon | 29 Feb 2024 | 60% |
| AI: Model Adaption (Stage 1) | Yeung Ho Yin Tommy,  Wong Ping Kuen | 15 March 2024 | 0% |
| AI: Model Training (Stage 2) | Yeung Ho Yin Tommy,  Wong Ping Kuen | 15 March 2024 | 50% |
| AI: Models Evaluation (Stage 3) | Yeung Ho Yin Tommy,  Wong Ping Kuen | 15 March 2024 | 50% |
| Web Application: UI design (Stage 1) | Li Chi Fung,  Rai Jon,  Yeung Ho Yin Tommy | 30 March 2024 | 60% |
| Web Application: Web Server Setup (Stage 3) | Yeung Ho Yin Tommy,  Wong Ping Kuen | 20 Jan 2024 | 100% |
| Web Application: Data Analysis (Stage 5) | Yeung Ho Yin Tommy,  Wong Ping Kuen | 1 Apr 2024 | 70% |

Key Successes and Failures

Successes:

* Completed the script for scraping OpenRice restaurant name and its review data.
* Completed the script for cleansing the scraped restaurant review data.
* Completed the model training.
* Completed the User Interface design for index page and result page.
* Completed the model deployment to the web application.
* Completed the script of the statistic function for the counting model’s outputs.
* Completed the statistic charts in the web user interface.
* Converted the dataset format from AAGCN to AI Challenger for model evaluation, it converted about 200 data.

Failures:

* I tried to use the Python module called “Scrapy” to implement the web scraping script, but I could not find a way to import the script into the web application, so I changed to use the Python module called “Beautiful Soup” to implement the web scraping.
* I tried to deploy the models of AAGCN and SSEGCN have the deployment issues that cannot accept the preprocessed data from datasets, this issue was found during the deployment steps, and we tried to fix this problem, but the problem is still not fixed.
* Although I created a restaurant review dataset with 300 data, but the dataset is not suitable for our AI Challenger model due to the dataset creation based on the AAGCN model, so that the created dataset cannot directly be used in AI Challenger model.
* I tried to write the script to connect Poe’s ChatGPT 3.5 service and it can send the message to obtain the ACSA prediction results. However, the Poe has the limitation that the free account only sends 100 messages to the ChatGPT 3.5 model every day.

## Appendix C2. (Group Project) Team Member #2's Interim Report

Wong ping Kuen (13031493)

Project Team Name: 2023-Keith-4

Project Title: Aspect-Based Sentiment Analysis for Online Restaurant Review in Chinese

Supervisor: Dr. Keith Lee

I, (Wong Ping Kuen, 13031493), certify that the description and information included in this team member's report is true to the best of my knowledge. In October 2023 and November 2023, I try to use “AAGCN” and “SSEGCN” to train the model and apply the model and custom dataset. In December 2023, I collected 300 restaurant review data for the dataset and used the ChatGPT to classify results for labeling each restaurant review. In January 2024, since “AAGCN” and “SSEGCN” couldn’t deploy to the web application, I made the Poe ChatGPT to analyze the restaurant review’ function more efficiently. Poe has a tokenizer limit and request time limit, so we found another ABSA model that supports Chinese and can be applied to the web server. Also, I made a script to evaluate the ABSA model’s accuracy and f1 score.

Wong ping Kuen, 21 January 2024

Task

|  |  |  |  |
| --- | --- | --- | --- |
| Tasks | Responsible Member(s) | Target Date | Completion |
| Technology Test: AI model | Yeung Ho Yin Tommy,  Wong Ping Kuen,  LI Chi Fung,  Jon Rai | 10 Jan 2024 | 100% |
| Datasets creation: data collection and data cleansing (Stage 1) | Yeung Ho Yin Tommy,  Wong Ping Kuen,  LI Chi Fung,  Rai Jon | 29 Feb 2024 | 60% |
| Datasets creation: Label the aspects for each review (Stage 2) | Yeung Ho Yin Tommy,  Wong Ping Kuen,  LI Chi Fung,  Rai Jon | 29 Feb 2024 | 60% |
| AI: Model Adaption (Stage 1) | Yeung Ho Yin Tommy,  Wong Ping Kuen | 15 March 2024 | 0% |
| AI: Model Training (Stage 2) | Yeung Ho Yin Tommy,  Wong Ping Kuen | 15 March 2024 | 50% |
| AI: Models Evaluation (Stage 3) | Yeung Ho Yin Tommy,  Wong Ping Kuen | 15 March 2024 | 50% |
| Web Application: Web Server Setup (Stage 3) | Yeung Ho Yin Tommy,  Wong Ping Kuen | 20 Jan 2024 | 100% |
| Web Application: Data Analysis (Stage 5) | Yeung Ho Yin Tommy,  Wong Ping Kuen | 1 Apr 2024 | 70% |

Key Successes and Failures

Successes:

* Completed the script for scraping and cleansing the OpenRice data that searched restaurant results.
* Completed the model testing.
* Completed the model training.
* Completed the script for evaluation of the model for prediction accuracy.

Failures:

* The models of AAGCN and SSEGCN have deployment issues that cannot train the model where the dataset is preprocessed.
* The Poe ChatGPT can make it more efficient, but that has tokenizer limitations and request limitations where we cannot fix this problem.
* Although I collected 300 restaurant review data, the dataset’ format is XML for “AAGCN” to the training model. Due to the dataset’s format not CSV that cannot directly be used in the new ABSA model.

## Appendix C3. (Group Project) Team Member #3's Interim Report

1. Information Box

Name of Student (and HKMU ID): LI Chi Fung 13031837

Project Team Name: 2023-Keith-4

Project Title: Aspect-Based Sentiment Analysis for Online Restaurant Review in Chinese

Supervisor's Name: Dr. Keith Lee

1. Declaration Statement

I, LI Chi Fung 13031837, certify that the description and information included in this team member's report is true to the best of my knowledge. In October 2023 and November 2023, I'm finding some research the Aspect-Based Sentiment Analysis and try some “ABSA” or “ACSA” code from the GitHub which related to “ABSA” and “ACSA”, also help to do the proposal report. In December 2023, I help to collect the dataset from openrice(300) and find a useful GitHub code about the aspect-based sentiment analysis which is called AI Challenger Sentiment Analysis. In January 2024, I help to design and develop the UI design. Moreover, I’m trying to find some model to improve the Ai challenger sentiment analysis which is using the BERT or using some online support system (AWS Amazon SageMaker).

LI CHI Fung 21/1/2024

1. Tasks Assigned to the Author and their Status

|  |  |  |  |
| --- | --- | --- | --- |
| Tasks | Responsible Member(s) | Target Date | Completion |
| Technology Test: AI model | Yeung Ho Yin Tommy,  Wong Ping Kuen,  LI Chi Fung,  Jon Rai | 10 Jan 2024 | 100% |
| Datasets creation: data collection and data cleansing (Stage 1) | Yeung Ho Yin Tommy,  Wong Ping Kuen,  LI Chi Fung,  Rai Jon | 29 Feb 2024 | 60% |
| Datasets creation: Label the aspects for each review (Stage 2) | Yeung Ho Yin Tommy,  Wong Ping Kuen,  LI Chi Fung,  Rai Jon | 29 Feb 2024 | 60% |
| Web Application: UI design (Stage 1) | Li Chi Fung,  Rai Jon,  Yeung Ho Yin Tommy | 15 Mar 2024 | 60% |
| Web Application: UI implementation (Stage 2) | Li Chi Fung  Rai Jon | 15 Mar 2024 | 60% |
| Survey design | LI Chi Fung  Rai Jon | 20 Jan 2024 | 50% |

1. Key Successes and Failures

Successes:

* Create the UI interface for the website
* find a model can use to analyze the Chinese comment.
* Try to find some model to improve the AI Challenger Sentiment Analysis.

Failures:

* I can’t help to create a graph displaying the results of the analyze comment and can’t change the Doughnut to display multiple data such as positive, neutral, and negative.
* I can’t help in backend develop, as for me the backend develop is difficult for now

## Appendix C4. (Group Project) Team Member #4's Interim Report

1. Information Box

Name of Student (and HKMU ID): Rai Jon 12749417

Project Team Name: 2023-Keith-4

Project Title: Aspect-Based Sentiment Analysis for Online Restaurant Review in Chinese

Supervisor's Name: Dr. Keith Lee

1. Declaration Statement

I, Rai Jon 12749417, certify that the description and information included in this team member's report are true to the best of my knowledge.

In October 2023 to November 2023, I have done online paper research about Aspect-Based Sentiment Analysis and in January 2024 we found the suitable model that directly support Chinese. I have test run the model with the web application.

At the end of December 2023, I worked on creating 300 datasets that were from OpenRice's review as the model's input and label the classification result for data by using ChatGPT to obtain the suitable result.

On During stage 1 and stage 2 of this project I have worked on Web Application building website structure including function that data retrieve from model is in sentiment polarity "-2, -1, 0, 1" then convert it into "1 as Positive", "0 as Neutral", "-1 as Negative" and "-2 as Not mentioned".

On January 2024, I created Survey for user feedback about the web application so we can improve.

Rai Jon, 24 January 2024

1. Tasks Assigned to the Author and their Status

|  |  |  |  |
| --- | --- | --- | --- |
| Tasks | Responsible Member(s) | Target Date | Completion |
| Technology Test: AI model | Yeung Ho Yin Tommy,  Wong Ping Kuen,  LI Chi Fung,  Jon Rai | 10 Jan 2024 | 100% |
| Datasets creation: data collection and data cleansing (Stage 1) | Yeung Ho Yin Tommy,  Wong Ping Kuen,  LI Chi Fung,  Rai Jon | 29 Feb 2024 | 60% |
| Datasets creation: Label the aspects for each review (Stage 2) | Yeung Ho Yin Tommy,  Wong Ping Kuen,  LI Chi Fung,  Rai Jon | 29 Feb 2024 | 60% |
| Web Application: UI design (Stage 1) | Li Chi Fung,  Rai Jon,  Yeung Ho Yin Tommy | 30 Mar 2024 | 60% |
| Web Application: UI implementation (Stage 2) | Li Chi Fung  Rai Jon | 15 Mar 2024 | 60% |
| Survey design | LI Chi Fung  Rai Jon | 20 Jan 2024 | 50% |

Successes:

* Completed web application structure and user interface(index).
* Completed scraping and cleansing the OpenRice data.
* Completed script for converting sentiment polarity into suitable classification on web application.

Failures:

* The dataset we created was not suitable for our current model because the dataset was created based on the previous model, AAGCN so the dataset was unable to be used on the current model.
* During the initial phase of this project, my contribution to the backend was limited due to my lack of experience in the field of AI. Additionally, some of the GitHub models were written in Chinese, which posed a challenge for me as an English speaker. As a result, I had to rely on the guidance of my groupmates, which led to slower project progress.

## Appendix D. (Group Project) Updated Meeting Minutes

|  |
| --- |
| Meeting date: 11 September 2023  Meeting time: 1 hour  Place: HKMU JCC  Attended members: All  Decisions made: Discussed and confirmed the project is topic ACSA |

|  |
| --- |
| Meeting date: 18 September 2023  Meeting time: 1 hour  Place: HKMU JCC  Attended members: All  Decisions made: Shared and tested the existing models |

|  |
| --- |
| Meeting date: 25 September 2023  Meeting time: 1 hour  Place: HKMU JCC  Attended members: All  Decisions made: Shared the ACSA findings |

|  |
| --- |
| Meeting date: 4 October 2023  Meeting time: 1 hour  Place: HKMU JCC  Attended members: All  Decisions made: Discussed the project proposal |

|  |
| --- |
| Meeting date: 16 October 2023  Meeting time: 1 hour  Place: HKMU JCC  Attended members: All  Decisions made: Discussed the initial report |

|  |
| --- |
| Meeting date: 1 November 2023  Meeting time: 1 hour  Place: HKMU JCC  Attended members: All  Decisions made: Discussed the initial report |

|  |
| --- |
| Meeting date: 6 November 2023  Meeting time: 1 hour  Place: Online  Attended members: All  Decisions made: Shared the finding about another existing ABSA model and existing Chinese dataset (AI Challenger) |

|  |
| --- |
| Meeting date: 4 December 2023  Meeting time: 1 hour  Place: Online  Attended members: All  Decisions made: Discussed the method of dataset creation and decided to use ChatGPT 3.5 model to classified and predict the aspects’ polarities. |

|  |
| --- |
| Meeting date: 9 January 2024  Meeting time: 1 hour  Place: Online  Attended members: All  Decisions made: Discussed the problems of AAGCN, SSEGCN, and Poe’s ChatGPT 3.5 model. We found the new existing ABSA model that directly support Chinese dataset and decided to use this model to implement our project. |

|  |
| --- |
| Meeting date: 23 January 2024  Meeting time: 1 hour  Place: HKMU JCC  Attended members: All  Decisions made: Reported the development progress of our prototype web application to Dr Keith. Decided to adapt the invalid format dataset to evaluate our AI Challenger model’s performance. |