Computer Science

COMP S456F - Software System Development Project

Final 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 : 21-04-2024

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 time collecting opinions from restaurant review websites, obtain more accurate analyzed review results, and also, to handle the lack of ABSA research in Cantonese. We labelled Cantonese restaurant review datasets for the model training and evaluation to create Cantonese Models. Also, we collected some Simplified Chinese datasets and applied the Microsoft Bing Translator to translate the Simplified Chinese datasets to Cantonese (Traditional Chinese). Then, we used the test dataset we created to evaluate the performance with 4 models and ChatGPT 3.5 turbo model for choose which model to be used. We created a web application that is used to deploy the trained ABSA model for analyze the OpenRice Chinese restaurant review. Lastly, we designed the user satisfaction survey to obtain the opinions from 15 test users for web application improvement.

Acknowledgement

We express our sincere gratitude to Dr. Keith Lee for his invaluable guidance and supervision throughout the entire process of our Final Year Project and he has spent his time on us. Also, he gave us a lot of materials about Deep Learning to help us to implement this project. It is thanks to his expert advice and insightful suggestions that we were able to gain a comprehensive understanding of the project and effectively approach its execution. We would like to extend our heartfelt appreciation to Dr. Roy Li for his valuable feedback and guidance during the project demonstration. His insightful advice played a pivotal role in helping us recognize the areas where our project fell short. This enabled us to identify and address the deficiencies, ultimately contributing to the improvement of our project.

Contents

1. Problem Definition

1.1. 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.

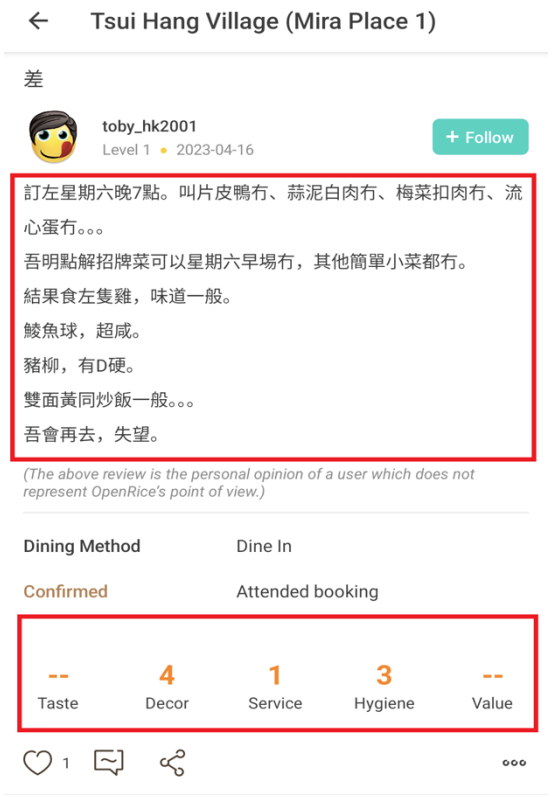


Figure 1: A user comment screen captured on OpenRice [1], the upper rectangle shows the reviewer's review about the restaurant, and the lower rectangle shows the rating for each aspect in restaurant by reviewer.

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 et al., 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.

[1] https://www.openrice.com/en/hongkong/review/%E5%B7%AE-e4787429

1.2. Finalized Project Aim

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.

1.3. Finalized Project Objectives

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 Cantonese 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.), and 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 Cantonese dataset.

- To train the existing models of Aspect-Based Sentiment Analysis (ABSA) with Cantonese datasets, and to evaluate the model's performance

- To develop a 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.

1.4. 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.

2. Literature Review

2.1. Related terminologies of Deep Learning

2.1.1. Accuracy

According to (Czakon, 2023), the results of the 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.

Here is the equation for Accuracy:

$$Acc = \frac{tp + tn}{tp + fp + tn + fn}$$

2.1.2. F1-Score

According to (Czakon, 2023), the 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.

Here is the equation for F1-Score:

$$\mathrm{F}_{beta} = (1 + \beta^2)\frac{precision \times recall}{\beta^2 \times precision \times recall}$$

2.1.3. Precision

According to (SerielWang, 2020), the 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.

Here is the equation for Precision:

$$Precision = \frac{TP}{TP + FP}$$

2.1.4. Recall

According to (SerielWang, 2020), the 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.

Here is the equation of Recall:

$$Recall = \frac{TP}{TP + FN}$$

2.2. Review of Existing or Related Solutions for the Problem

2.2.1. 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).

2.2.2. 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 et al., 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** aspect and **"environment"** to **Hygiene** aspect.

According to (Li et al., 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.

2.2.3. Existing Models for ACSA and ABSA

According to (Liang et al., 2018), 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 et al. 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

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 for the development of AI technologies. Sentiment Analysis is one of the 2018 challenges in AI challenger. According to (JohanyCheung, 2018), **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 Simplified or Transitional Chinese data since that used "OpenCC" model translates data into Simplified or Transitional Chinese.

The limitation of the model is that mostly supports the prediction of the Simplified Chinese data since the model's preprocessing part handles the Simplified Chinese data. Therefore, the Cantonese data cannot be handled very well in this model.

2.2.4. Neural Network used in the existing models

According to (O'Shea & Nash, 2015), **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.

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 et al., 2014), and Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering (Defferrard et al., 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.

2.2.5. 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.

2.2.6. 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.

2.3. 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 the adapted AI Challenger Transformer+Convolutional model
Support Chinese Dataset	No	No	Yes (only Simplified Chinese)	Yes (only Traditional Chinese (Cantonese))
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	85.38	64.59
Our Solution	AI Challenger Sentiment Analysis Training set 2018 (45,000 translate to Cantonese) + Our Cantonese Training Dataset (500 Cantonese) Our Cantonese Evaluation Dataset (1,000 Cantonese)	78.67 (final result)	44.76 (final result)

Table 2: Solution Performance Comparison

3. Methodology

3.1. 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 OpenRice 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 Cantonese dataset. Furthermore, we have collected, cleansed, and labelled 1500 Chinese (Cantonese) restaurant reviews from OpenRice manually, also, collected the AI Challenger Sentiment Analysis Training set 2018 from the internet and we found 2 methods to translate the data from Simplified Chinese to Cantonese. To evaluate the performance of our web application, we used the created Cantonese restaurant dataset to evaluate our model’s performance, also, applied the prompt technique to obtain the Cantonese restaurant dataset’s classified and labelled results from the ChatGPT 3.5 model, to evaluate its performance in handle ABSA task with a Cantonese restaurant review. Furthermore, we created a user satisfaction survey to obtain user feedback to improve the user experience of our web application.

3.2. Requirements and Key Technologies

3.2.1. 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

Table 3: Function List

3.2.2. 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.

Table 4: Functional requirements

3.2.3. 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, and we also use the Font Awesome icon library to make the user interface more user-friendly ("Use Font Awesome on the Web," n.d.). 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). 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 use. Moreover, we will use the OpenCC module (BYVoid, 2023) to translate the Simplified Chinese data web application retrieved data to Cantonese, also, we are using Microsoft Bing Translate API ("Text Translation," n.d.) to help us translate the AI challenger dataset from Simplified Chinese to Cantonese. 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, also, we will use the ChatGPT 3.5 model combined with the ChatGPT prompt technique (Su, 2023) to help us to predict the test dataset and evaluate its predicted results performance by scikit-learn.

3.2.4. 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 classify the aspects and its polarities, and add the result (review data and classified results) to the Cantonese dataset (training and evaluation). Uses the Microsoft Bing Translator API to translate the review data of AI Challenger dataset from Simplified Chinese to Chinese (Cantonese).
- Stage 3: Adapt the code of the AI Challenger model to handle our created Cantonese (Traditional 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. Moreover, to create a ChatGPT prompt combined with reviews from evaluation dataset to predict the aspects' sentiment, and evaluate it prediction quality by use the scikit-learn module to compare the original evaluation dataset with the predicted result.

- 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 data that potentially occurred Simplified Chinese text to Cantonese as the model's input.

- Stage 7: Use Bootstrap 5 front-end framework, Font Awesome icon library, 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.

3.3. System Design

3.3.1. 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 for restaurant improvement.

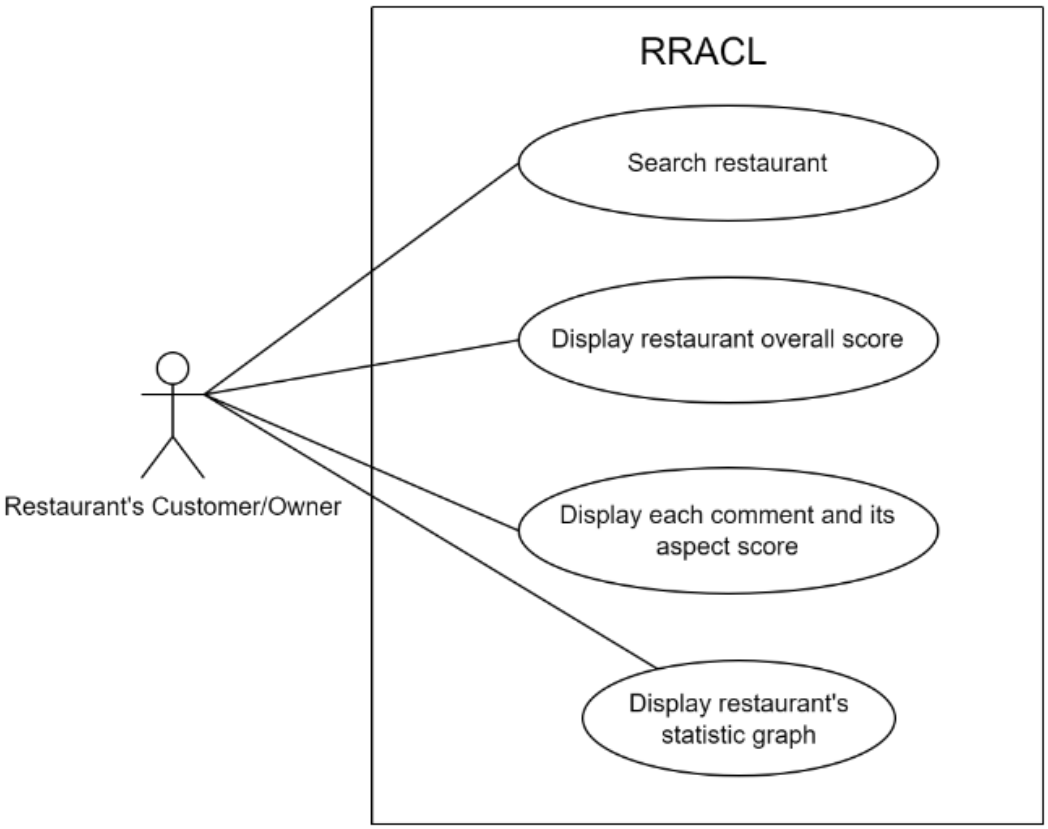


Figure 2: Use Case of the system Restaurant Review Analyzer for Chinese Language (RRACL)

3.3.2. Architecture or High-Level Design

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

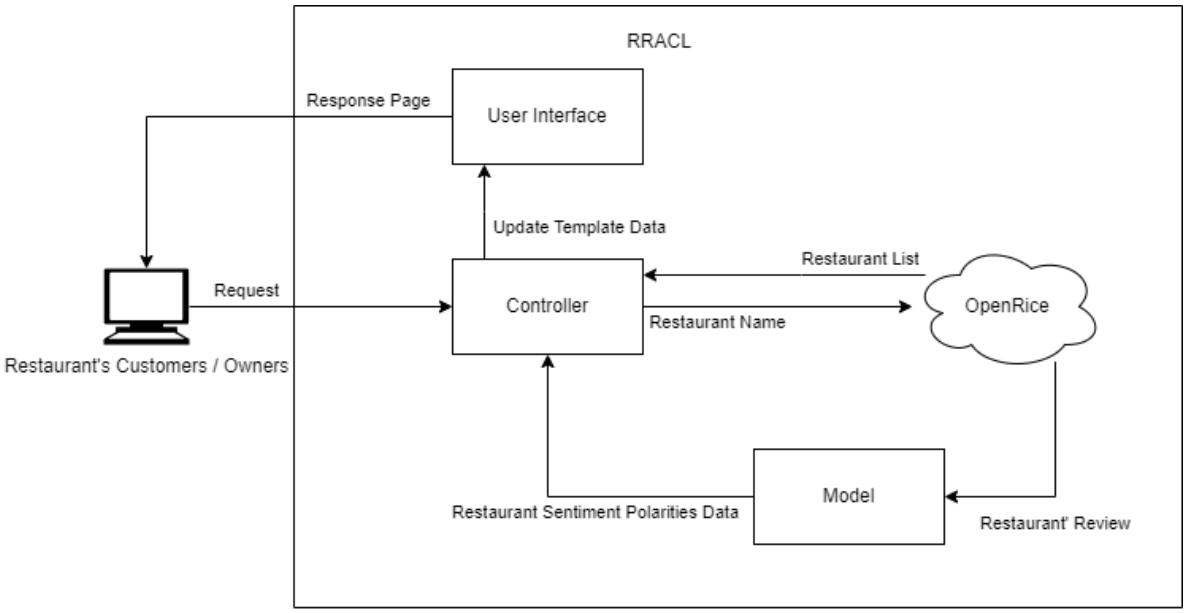


Figure 3: High-Level System Design of the system Restaurant Review Analyzer for Chinese Language (RRACL)

- **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.

3.3.2.1. Component diagram

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

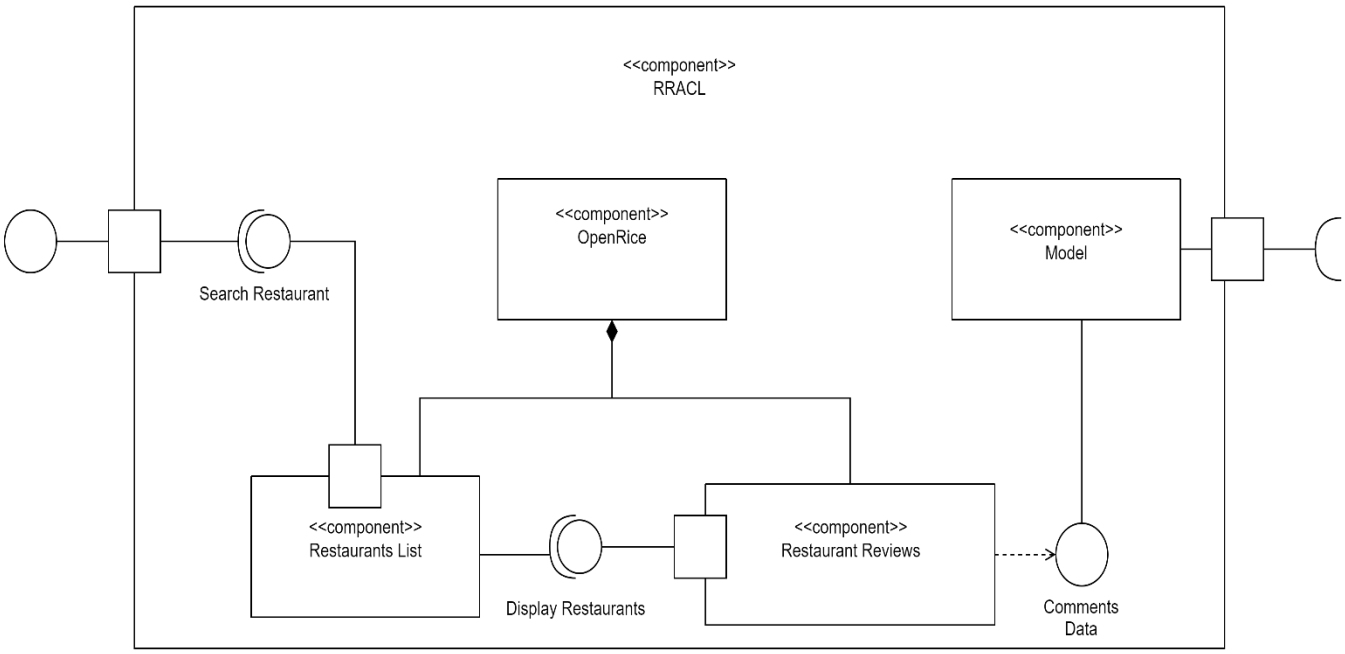


Figure 4: Component of the system Restaurant Review Analyzer for Chinese Language (RRACL)

- **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.

3.3.2.2. Data-flow diagram

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

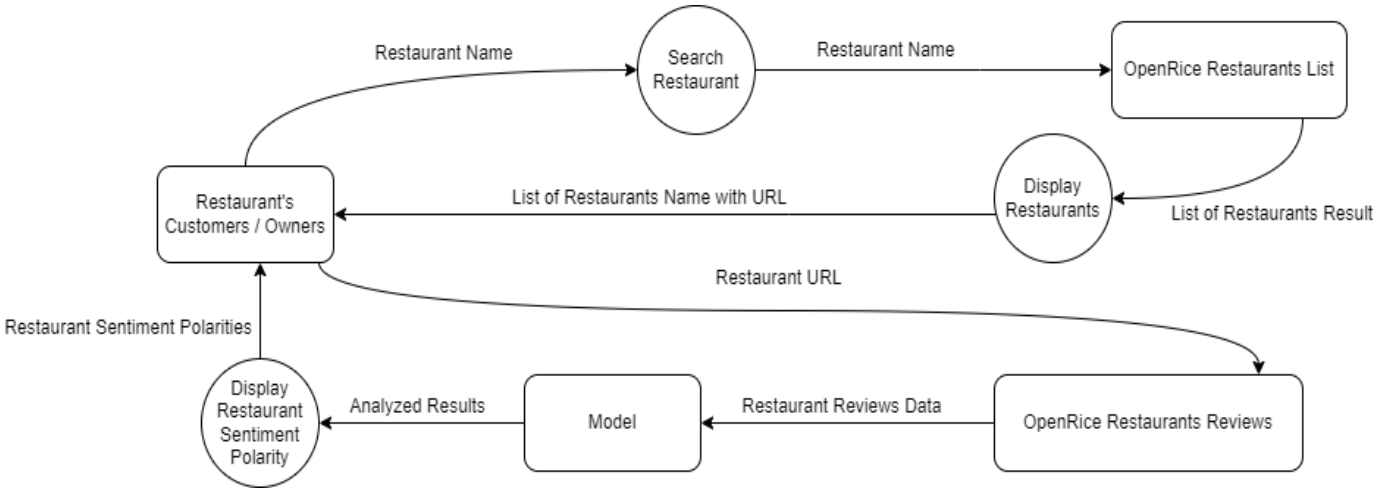


Figure 5: Data-flow of the system Restaurant Review Analyzer for Chinese Language (RRACL)

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

3.3.3. Sequence diagram

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

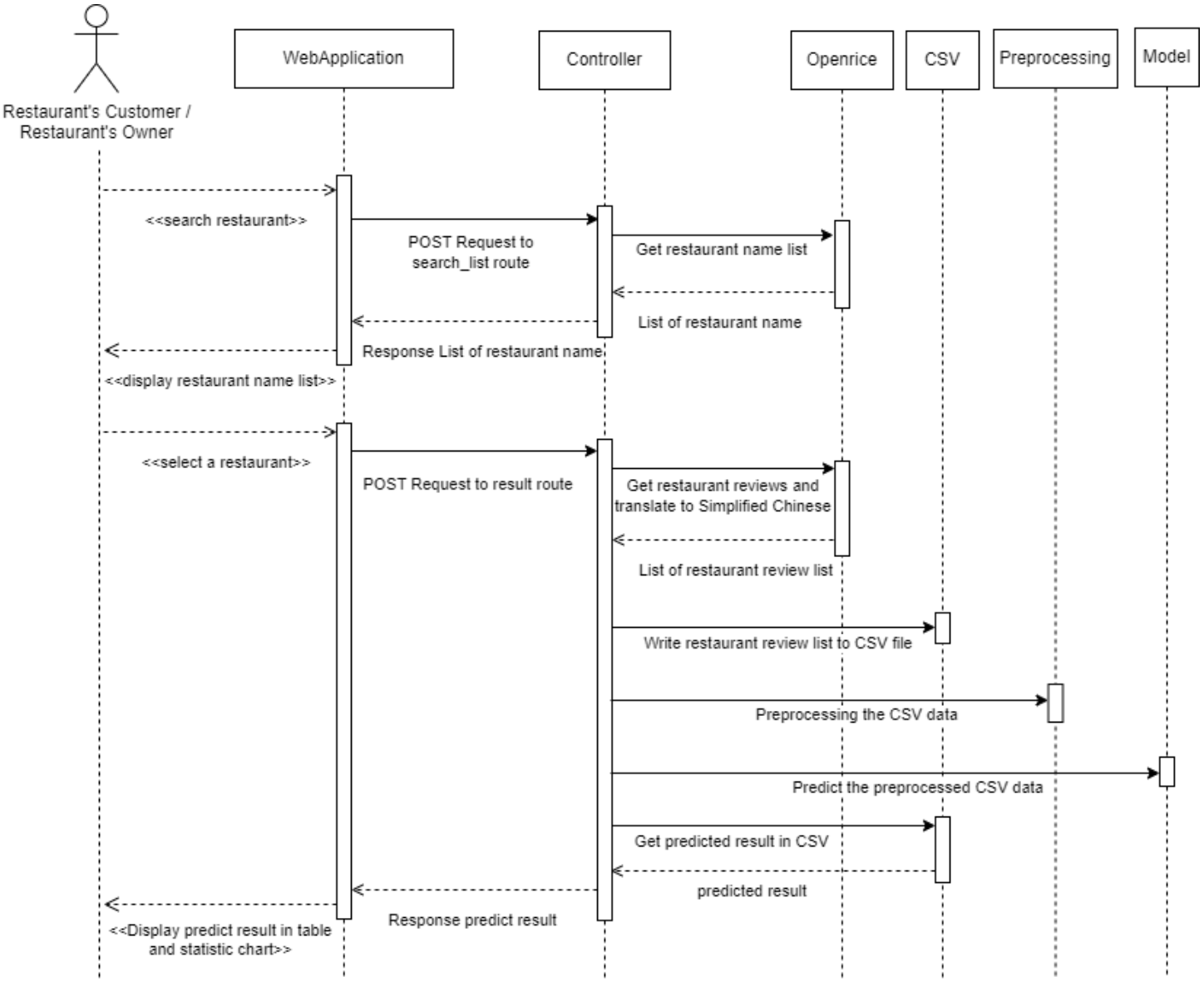


Figure 6: Sequence diagram of the system Restaurant Review Analyzer for Chinese Language (RRACL)

3.4. Dataset Collection and Creation

For the model training and evaluation, we collect the AI Challenger Sentiment Analysis dataset from the internet. Furthermore, as the Cantonese restaurant reviews is lack of dataset, therefore, we are manually to create our own dataset by collects the restaurant reviews in OpenRice and follow the label of AI Challenger Sentiment Analysis dataset to labelling the data. In the AI Challenger Sentiment Analysis Training set 2018, most of the data are Simplified Chinese and written language. That means not all datasets are in line with our system that the model training dataset. Therefore, we write a script automatically to use the "Microsoft Bing Translator API" for translate the Simplified Chinese review data to Cantonese.

3.5. Evaluation Method

In this project, we implement two type of evaluation method which are performance evaluation and user evaluation. For the performance evaluation, we use the collected training dataset to train 4 models and use the 1 evaluation dataset to evaluate 5 different models' performance in the Cantonese restaurant review in Aspect-Based Sentiment Analysis, including Simplified AI Challenger Models, Cantonese AI Challenger Models, and the ChatGPT 3.5 turbo Model. Moreover, we will design and initiate the User Satisfaction Survey for the test users of our system to collect their opinions. This will find 15 test users to use our system and fill out the survey.

4. Results – Prototype Implementation

4.1. Implementation Issue

For our prototype, we develop web application based on the high-level programming language Python 3.6.5. Then, we use the Beatifulsoup 4.12.2 to scraping the OpenRice restaurant and its review data. For the application, we deployed an Aspect-Based Sentiment Analysis (ABSA) model to perform the sentiment analysis. Also, we use OpenCC 1.1.1 module to translate the scraped review data to Cantonese 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, Font-Awesome icon library, and Chart.js 2.9.4 visualization libraries to create.

4.2. Prototype Operating and User Interface

At the final stage of the website interface design, we implemented numerous modifications to ensure that users can navigate the website with ease and without encountering any issues. To enhance the user experience on our website, we consciously incorporated more visual elements such as captivating images and icons sourced from the Font Awesome library. By doing so, we aimed to facilitate a more intuitive understanding of our content, allowing users to effortlessly comprehend our message at a glance.

To elevate the functionality of our website, we have skillfully integrated technologies such as Bootstrap 5, JavaScript, and Flask. This has enabled us to create dynamic user interfaces that are both visually appealing and highly interactive. Additionally, we have incorporated charts to effectively showcase the results of each aspect, providing users with a comprehensive overview of the restaurant's performance. Our website comprises of three key pages - the Main page, Search page, and Result page - each designed with meticulous attention to detail to ensure a seamless user experience.

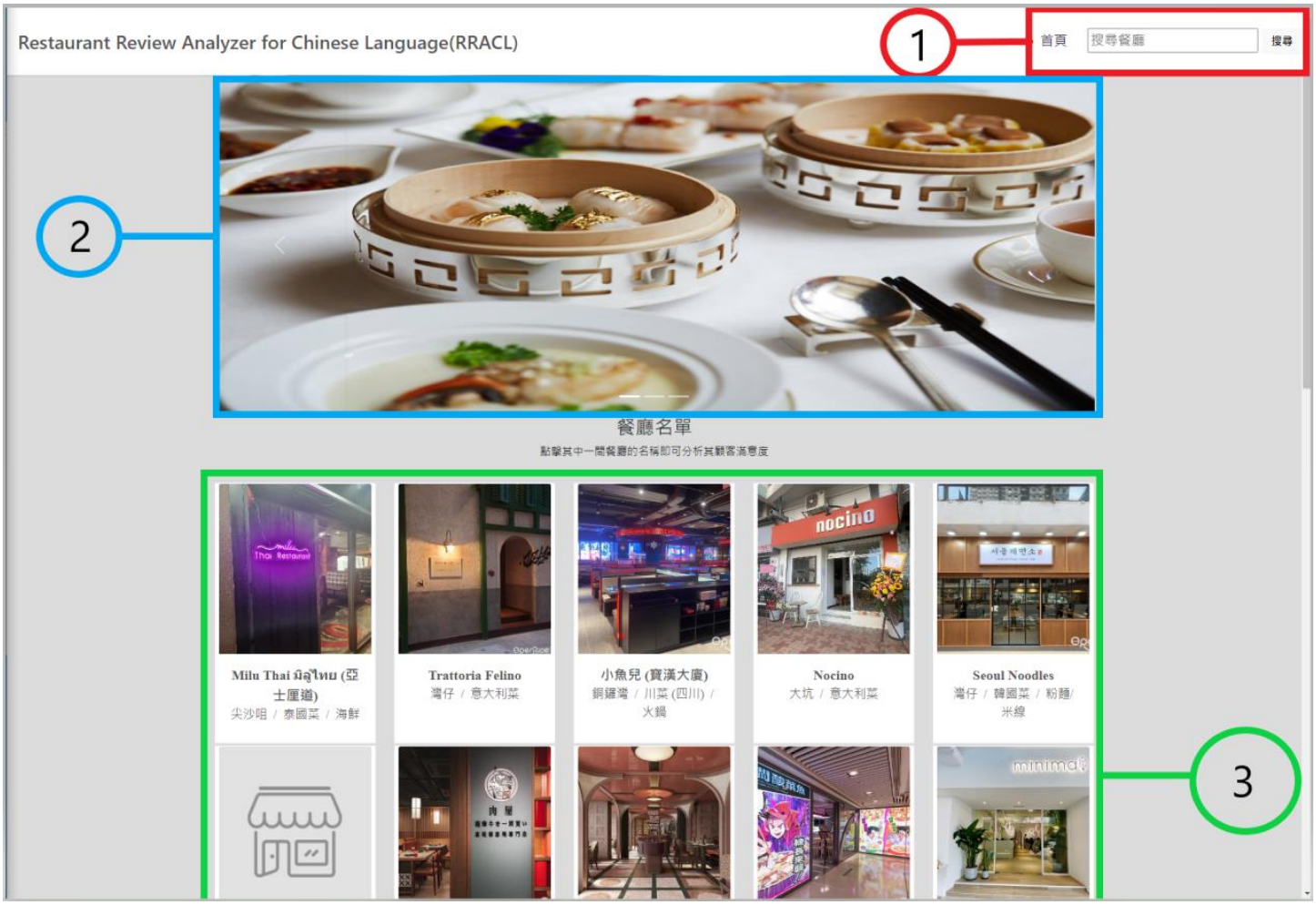


Figure 7: Main Page will show top 30 restaurant of Open rice.

For this page data, we use Python BeautifulSoup library to scrape the data from the OpenRice Restaurant Chart page, including the restaurant pictures, restaurant names, restaurant types, and the restaurant page URL. Moreover, we implement the search restaurant function by sending the request using the OpenRice URL.

1. Our website's main page have included a user-friendly search bar, conveniently located on the navigation bar as shown in Figure 7 allowing users to search for their desired restaurant by name. The user input will pass as a parameter of the OpenRice search URL to search for a list of restaurants in OpenRice.

2. To further enhance the user experience, we have also included three images of each restaurant, providing a glimpse of their exquisite ambiance and decor.

3. Additionally, a list of Open Rice's top 30 restaurants, elegantly displayed in Figure 7. This figure displays the data that was mentioned which are restaurant names, and restaurant pictures. These data scrape from the OpenRice.

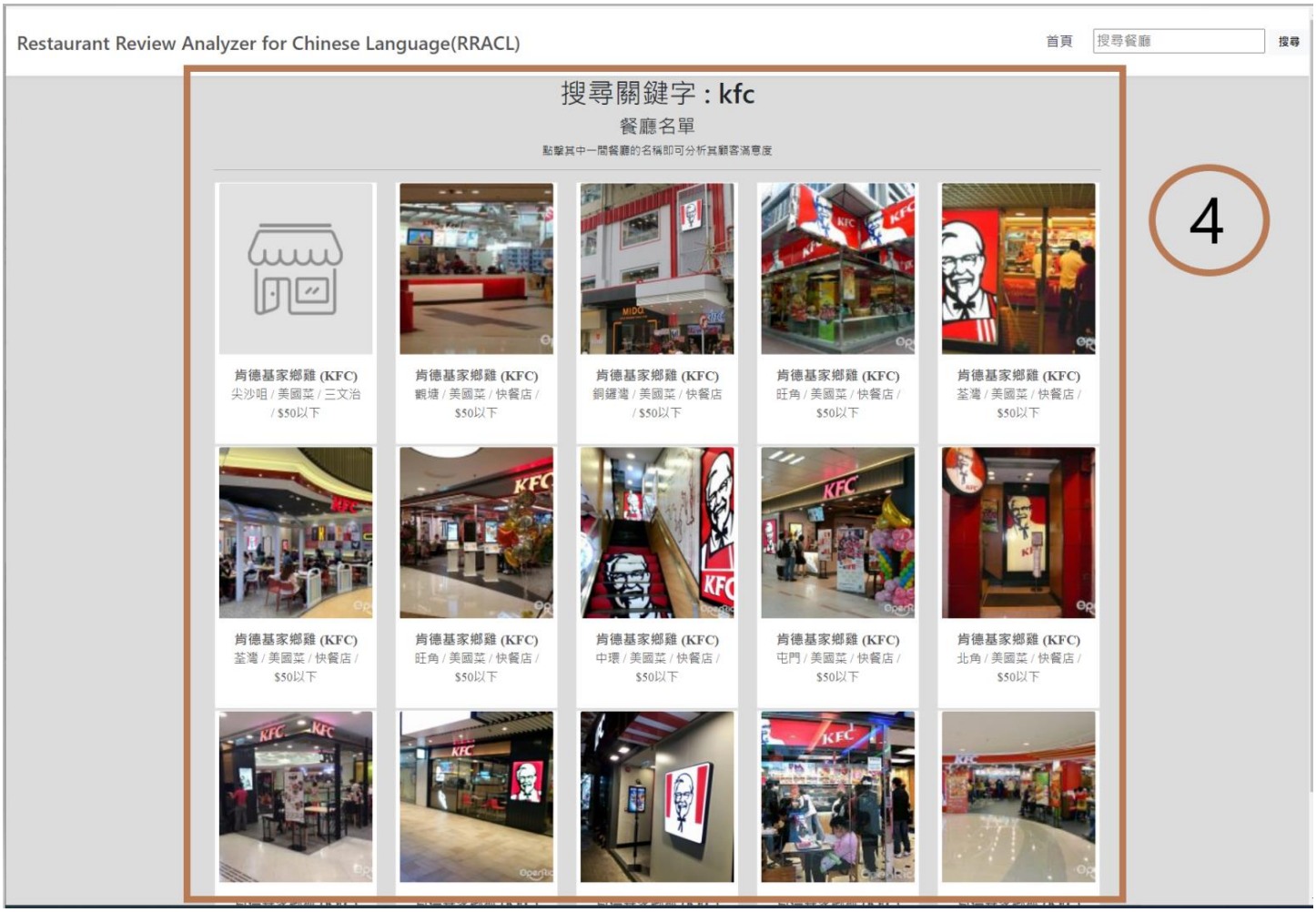


Figure 8: Search Page displaying result in restaurant detail.

Our Search Page, elegantly displayed in Figure 8, presents a comprehensive list of restaurants that match the user's search keyword through the search bar, these results are retrieved from the OpenRice search result page by the use of the BeautifulSoup module. Each restaurant is accompanied by a picture, name, district, and average spending, providing users with a detailed overview of their options. Our user-friendly interface ensures that users can easily navigate through the search results and find the perfect restaurant to suit their preferences.

4.) list of restaurants with picture, name district and average spending with the searched keyword. These data also use BeautifulSoup to scrape from search result page in OpenRice.

Figure 9: This is the loading animation when user have wait for restaurant analyze results.

In Figure 9, the loading animation will be presented when the user clicks on one of the restaurants. During the loading animation shown, the application will scrape the selected restaurant data from OpenRice, it includes the restaurant reviews and some restaurant-related information. The scraped restaurants review will use the OpenCC module translate text that is Simplified Chinese to Cantonese. Then, the web application will use pandas module to write the review data to a CSV file as the dataset and pass it to the model for prediction. After the prediction step is finished, it will display the restaurant-related information and use the pandas module to read its analyzed result obtained from the prediction result dataset, and then display it on the page as Figure 10 and Figure 11.

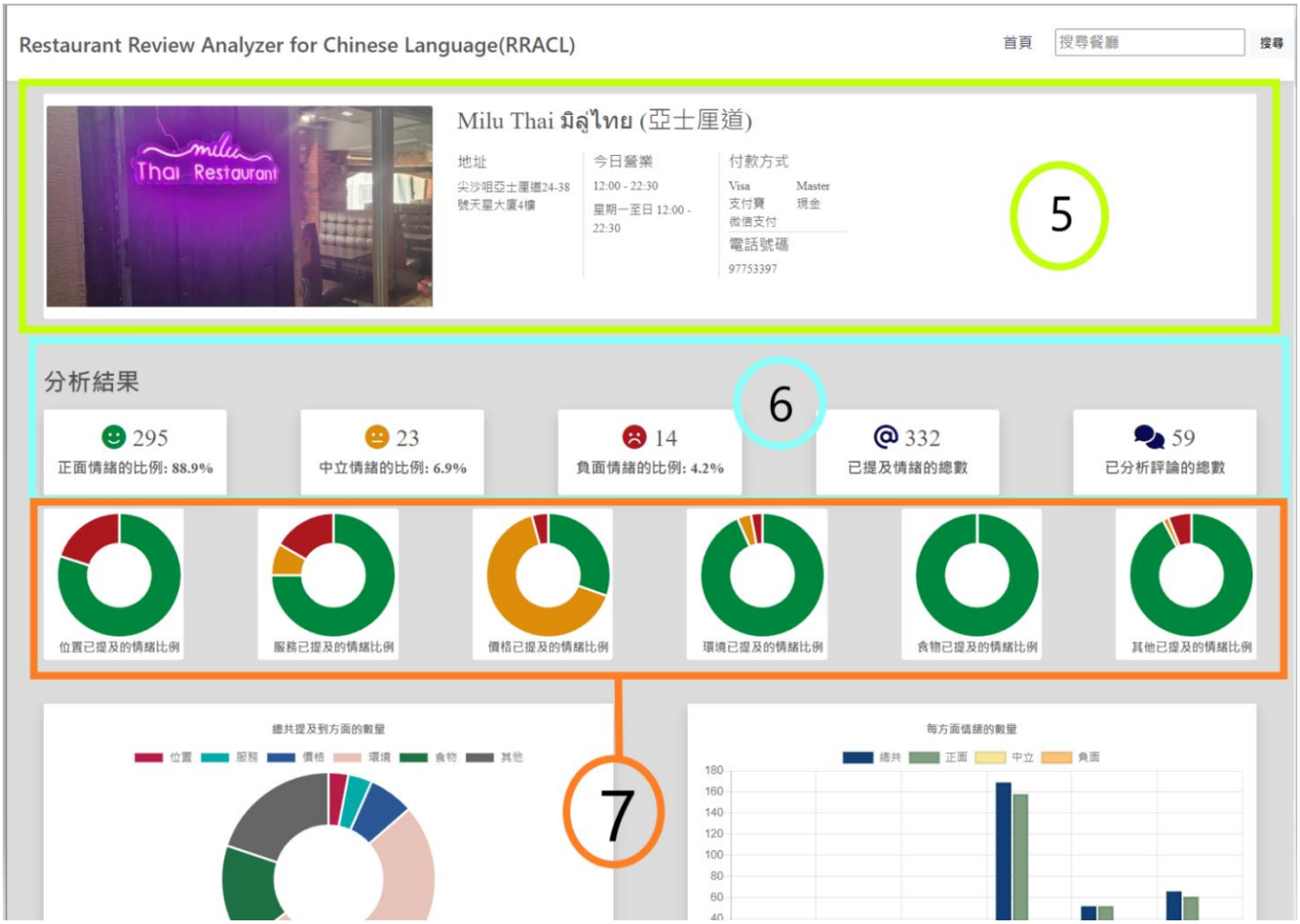


Figure 10: Result Page showing the restaurant's information and analyze results in graphs.

Our result page provides users with a comprehensive overview of each restaurant, including restaurant-related information such as the number of reviews and the total number of positive, neutral, and negative reviews. For example, 'Milu Thai' displays 59 reviews, with a total of 295 positive, 23 neutral, and 14 negative reviews, resulting in a total of 332 polarities. Additionally, six graphs are displayed, representing each category of positive, neutral, and negative reviews, providing users with a visual representation of the restaurant's overall rating. Our user interface ensures that users can easily access and analyze the information. The statistical results are based on the model output dataset that to calculate. These statistical charts are use the Chart.js visualization libraries to create. The icons in the (5) are using the Font-Awesome icon library.

5.) Shows the restaurant information e.g.(address, opening hour, telephone and payment)**.**

6.) From left to right, it shows the statistic results of **Total positive aspects and percentage, Total neutral aspects and percentage, Total negative aspects and percentage, Total number of mentioned polarities** and **Total Analyzed Reviews**:

Total number of mentioned polarities – it is the result of the total number of all mentioned polarities summed up, the mentioned polarities are counted from model outputs. Below is the calculation formula:

$$\boldsymbol{Total\ number\ of\ mentinoed\ polarities = Total\ positive\ numbers + Total\ neutral\ numbers + Total\ negative\ numbers}$$

Total Analyzed Comments – It is the number that represents the comments we collected from OpenRice and analyzed by the model.

Total positive aspects and percentage – This is the result shown that the total number of mentioned positive, and the total number of mentioned positive occupied percentage in the Total number of mentioned polarities. Below is the calculation formula:

$$\boldsymbol{Positive\ Polarities\ Percentage} = (\boldsymbol{Total\ number\ of\ positive} \div \boldsymbol{Total\ number\ of\ mentioned\ sentiment\ polarities}) \times \mathbf{100}$$

Total neutral aspects and percentage – This is the result shown that the total number of mentioned neutral, and the total number of mentioned neutral occupied percentage in the Total number of mentioned polarities. Below is the calculation formula:

$$\boldsymbol{Neutral\ Polarities\ Percentage} = (\boldsymbol{Total\ number\ of\ neutral} \div \boldsymbol{Total\ number\ of\ mentioned\ sentiment\ polarities}) \times \mathbf{100}$$

Total negative aspects and percentage – This is the result shown that the total number of mentioned negative, and the total number of mentioned negative occupied percentage in the Total number of mentioned polarities. Below is the calculation formula:

$$\boldsymbol{Negative\ Polarities\ Percentage} = (\boldsymbol{Total\ number\ of\ negative} \div \boldsymbol{Total\ number\ of\ mentioned\ sentiment\ polarities}) \times \mathbf{100}$$

7.) This part shows the doughnut charts of each aspect in Location, Service, Price, Environment, Dish and Others. Each doughnut chart 3 type of value. For example, the Location doughnut chart has the total number of location positive, total number of location neutral, and total number of location negative. These values also have its percentage, and the location aspect formula are listed below:

$$\boldsymbol{Location\ Positive\ Polarities\ Percentage} = (\boldsymbol{Total\ number\ of\ location\ positive} \div \boldsymbol{Total\ number\ of\ location\ sentiment\ polarities}) \times \mathbf{100}$$

$$\boldsymbol{Location\ Neutral\ Polarities\ Percentage} = (\boldsymbol{Total\ number\ of\ location\ neutral} \div \boldsymbol{Total\ number\ of\ location\ sentiment\ polarities}) \times \mathbf{100}$$

$$\boldsymbol{Location\ Negative\ Polarities\ Percentage} = (\boldsymbol{Total\ number\ of\ location\ negative} \div \boldsymbol{Total\ number\ of\ location\ sentiment\ polarities}) \times \mathbf{100}$$

Figure 11: Result Page showing the restaurant's analyzed result charts and comments.

8.) Two graphical representations, namely a doughnut chart and a bar chart, will be presented, showcasing the polarities of each aspect within the restaurant.

9.) The comments sourced from OpenRice are retrieved and subjected to analysis, resulting in the assignment of rating scores to each comment. Additionally, a button will be provided to access the corresponding bar chart to have more detail on the analysis result.

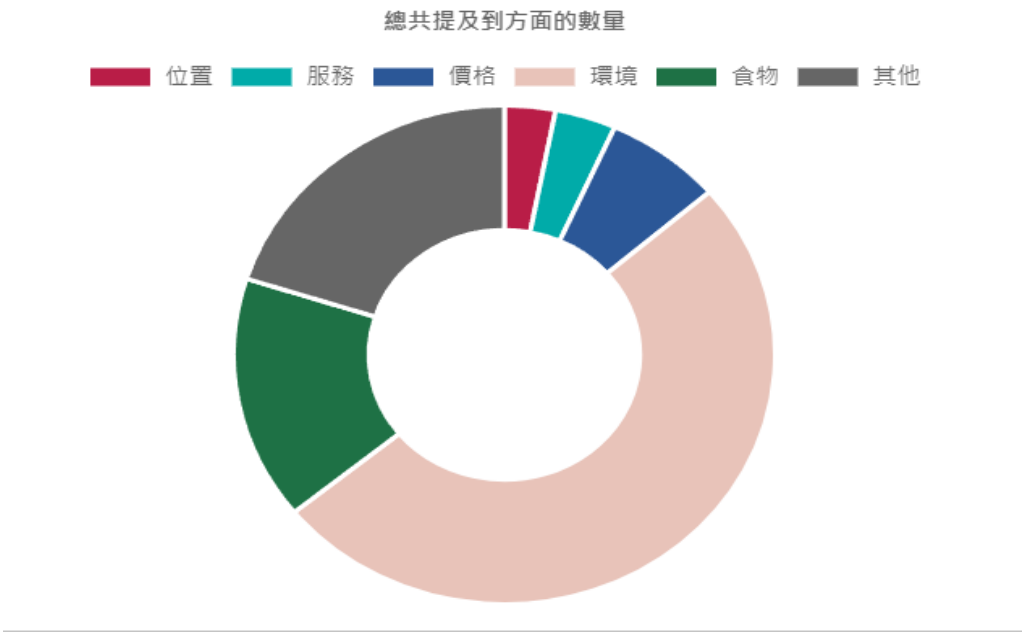


Figure 12: The doughnut chart of the total number of mentioned aspects

In figure 12, it will provide an overview of the overall polarities within each of the six categories, namely location, service, price, environment, food, and others. Each aspect in this chart sum up all its polarities which are positive, neutral, and negative sentiments.

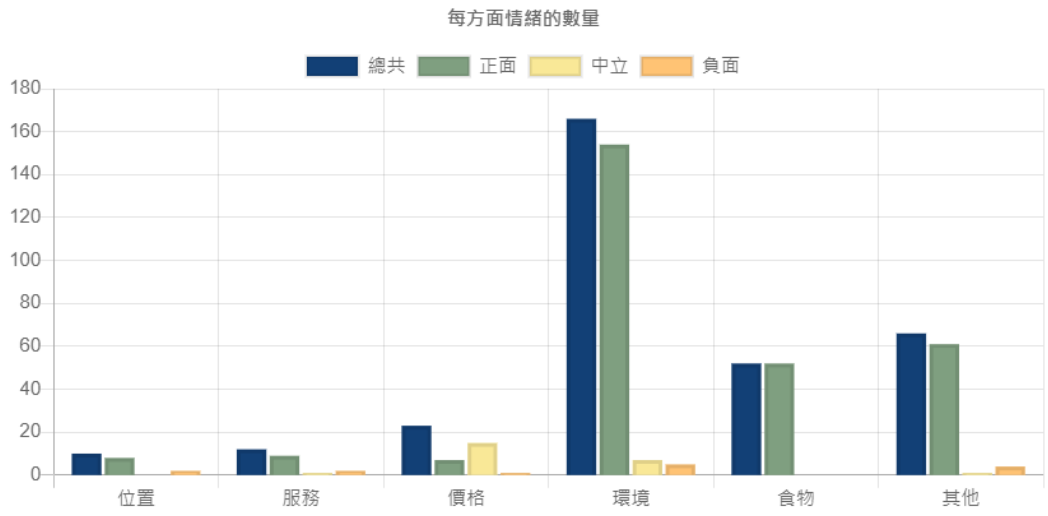


Figure 13: The bar chart of the total number of each aspect

This bar chart will illustrate the cumulative polarities for each aspect, including the total number of all sentiments, the total number of positive sentiments, the total number of neutral sentiments, and the total number of negative sentiments.

Figure 14: Result Page showing restaurant comments and rating scores.

10.) Represent as rating score by using stars (1 - 5).

11.) Upon clicking the "Detail" button, which represents the stacked bar chart, will be displayed.

Figure 14 and Figure 15 depict the rating scores and "Detail" buttons associated with each comment. Clicking the "Detail" button will reveal a dedicated stacked bar chart representing the polarity of the respective comment. This ensures that each comment is accompanied by its own rating score and a corresponding stacked bar chart for visual representation. The rating score is calculated by the mentioned polarities in each comment, the calculation formula and rules are listed below:

$$\textbf{\textit{Rating Score}} = \frac{(\textbf{\textit{Total number of positive in the comment}} - \textbf{\textit{Total number of negative in the comment}})}{(\textbf{\textit{Total number of positive in the comment}} + \textbf{\textit{Total number of neutral in the comment}} + \textbf{\textit{Total number of negative in the comment}})}$$

The stars display depends on the rating score result range, the score is result in -1 to 1. The range to display the stars are listed below:

- -1 display 0 star.
- -0.9 to -0.6 will display 1 star.
- -0.5 to -0.2 will display 2 stars.
- -0.1 to 0.2 will display 3 stars.
- 0.3 to 0.6 will display 4 stars.
- 0.7 to 1 will display 5 stars.

Figure 15: Upon clicking the "Detail" button, which represents the stacked bar chart, will be displayed.

12.) The stacked bar chart presents the outcomes of the individual comments in a visual format, showcasing the corresponding results.

The stacked bar chart will present the results for any mentioned sub-aspects that are under the aspect in the corresponding comment, specifically indicating positive, neutral, or negative outcomes. If a sub-aspect is not mentioned in the comment, it will be excluded from the stacked bar chart. For example, the location aspect has 3 sub-aspects which are location traffic convenience, location traffic convenience, and location easy to find, if there is only 1 sub-aspect is mentions its sentiment polarity like the location traffic convenience with positive sentiment, it will only display the sub-aspect location traffic convenience bar with positive sentiment in the chart.

5. Results – Evaluation

In this section, we will discuss how to adapt the existing Aspect-Based Sentiment Analysis model to handle Cantonese reviews, and also talk about how to collect the dataset and how to evaluate the performance with the adapted AI model. Moreover, we discuss how to evaluate our prototype effectiveness by initiating the user satisfaction survey. Lastly, we will discuss the evaluation results of the model performance and the prototype performance.

5.1. Model Structure

In this section, we will talk about the structure of AI Challenger model and the model adaption for handle the Cantonese restaurant review. For this model, it is a multi-task model that receive a review input and return output for 20 aspects, in Figure 16, it is training by two layers which are the preprocess layer and Convolutional Neural Networks (CNN) layer. For the preprocessing layer, it is used to preprocess the dataset for the next layer to train or predict, including the word segmentation, word embedding, position embedding and three layers transformer. For the CNN layer, it predicts the sentiment data to 20 aspects. The layers include convolution, max pooling and fully connected and softmax. This model is trained by the 4 types of labels which are 1 (positive), 0 (neutral), -1 (negative), -2 (not mentioned), and these labels are the Integers.

- Stop word filtering: This step uses the stop word from Weibo Simplified Chinese data to remove each sentence word where the word is the same in the stop word list.

- Word Segmentation: This step uses the 'jieba' library to segmentate sentences by 'jieba' default dictionary (Simplified Chinese).

- Word Embedding: This layer encodes the words that are closer in the vector space and are expected to be similar in meaning. The model uses word2vec trained with Weibo corpora for word embedding.

- Position Embedding: This layer sends the word position information (sequence) into the next layer (Transformer).

- Transformer: These layers use a self-attention encoder. It calculates the correlation score of each element with other elements where the input sequence interacts (Raganato et al., 2020).

- Convolution: This layer is a convolution performed between a sentence and a kernel equal to a sequence matrix.

- Max Pooling: This layer takes the maximum value within the sequence and the maximum value becomes a single sequence.

- Fully connected: This layer weighed linear transformation to the input neurons and then passed the output through a non-linear activation function, it returns the (model's output) logits and pass to the next softmax layer.

- Loss function: It uses the Categorical Cross Entropy Loss as the loss function. According to The 365 Team (2023) (What Is Cross-Entropy Loss Function?, 2023), Categorical Cross Entropy Loss used to calculate the different between the logits and the expected labels (the original labels) in the softmax layer, it used to minimize the different during the model training.

- Softmax: This layer will output each aspect sentiment result is assigned a probability value

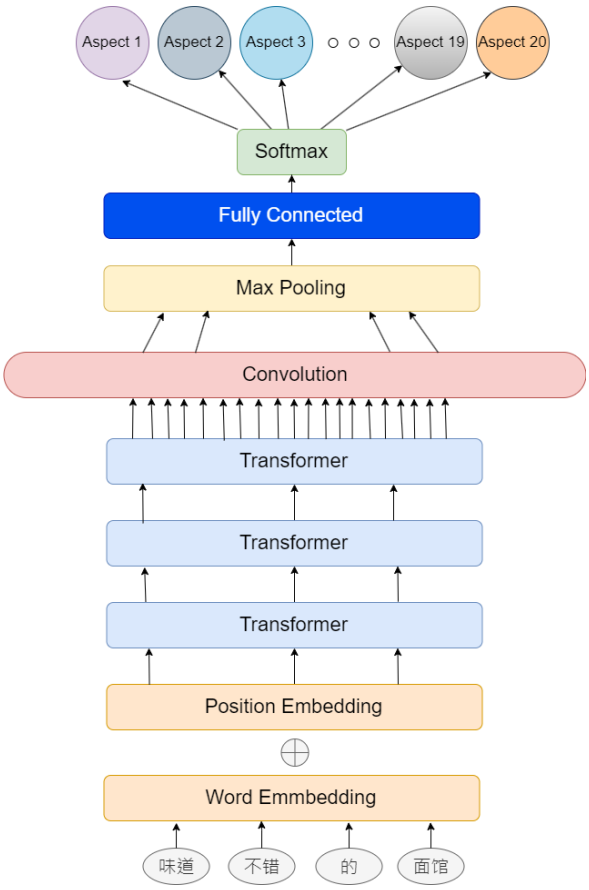


Figure 16: Structure of AI Challenger Model

5.2. Model Adaptation

To enhance the ability of AI Challenger model in processing Cantonese review, we adapt the partial preprocessing layer in the model which are the word segmentation and word embedding. For the word embedding, we use the fastText vectors created from ToastyNews data (ToastyNews & lso, 2020) as the word embedding in our model. We also translate the original stop word list translate to Traditional Chinese and expand the list by add the stop word provided by PyCantonese (Lee et al., 2022). For the word segmentation, we still use the jieba word segmentation library and load the external Cantonese dictionary to segment the Cantonese words (wchan757, 2019).

5.3. Model Evaluation Method

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 Macro-F1 score. The Accuracy score for each label is produced by setting the predicted data list and the 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.

5.4. Datasets

In this section, we are talk about the dataset that we used to train and evaluate the model, and the details are listed below:

- **Training datasets**: The **AI Challenger Sentiment Analysis Training set 2018** dataset is used in the original AI Challenger Model that contains 101,000 Simplified Chinese restaurant review data, and we only keep the first 45,000 data for model training. Furthermore, we also created a new dataset that use the Microsoft Bing translator API (Text Translation, n.d.) to translate those 45,000-review data to Traditional Chinese (Cantonese), and also extend this dataset by collect 500 Cantonese review data from OpenRice and labelled manually. Moreover, we also created a small dataset that uses the 500 review data from the same translated AI challenger dataset and append the 500 OpenRice review data as mentioned before. Therefore, there are three datasets for model training which are the **45,000 AIC dataset,** the **45,000 AIC + 500 OpenRice train dataset, and 500 AIC + 500 OpenRice train dataset**, the data statistic detail are listed on Table 6.

- **Evaluation datasets**: We also created a dataset that manually collected 1000 Cantonese restaurant reviews from OpenRice and labelled them. To evaluate the model that supported different languages, we created a new dataset with the same set of data that we only translated the review data from Cantonese to Simplified Chinese by the Microsoft Bing Translator API. Thus, there are two datasets which are the **1,000 OpenRice Evaluate dataset (Cantonese)** and the **1,000 OpenRice Evaluate dataset (Simplified Chinese)**, the data statistic detail are listed on Table 6.

Dataset	Short Terms	Division	Dataset size	Positive	Negative	Neutral	Not mentioned
45,000 AI Challenger Sentiment Analysis Training set 2018 (Original)	**45,000 AIC dataset**	Train	45,000 review data	220,409	38,361	77,913	563,317
500 translated AI Challenger training set 2018 + Our 500 training dataset (Cantonese) or (Translate to Simplified Chinese)	**500 AIC + 500 OpenRice train dataset**	Train	45,500 review data	222,413	38,522	78,066	570,999
45,000 translated AI Challenger training set 2018 + Our 500 training dataset	**45,000 AIC + 500 OpenRice train dataset**	Train	1000 review data	127	7	5	892
OpenRice evaluation dataset (Cantonese) or (Translate to Simplified Chinese)	**1,000 OpenRice Evaluate dataset**	Test	1,000 review data	3,986	1,041	771	14,201

Table 6: Statistics for the datasets

5.5. Experiment Setting

For our experiment, we have set up 5 different models to evaluate the performance between these five models. For Models 1 and 2, both models are Simplified Chinese model that use the 45,000 AIC dataset and 500 AIC + 500 OpenRice (Simplified Chinese) to train dataset to train the model 1 and model 2 respectively. For the embedding, we still use the word2vec which is trained with Weibo corpora and the positional embedding, it uses the Momentum Optimizer with a learning rate of 0.01, and a batch size of 32. This optimizer sets the dropout rate to 0.1. For Models 3 and 4, both models are Cantonese models that also use the 45,000 AIC + 500 OpenRice train dataset and 500 AIC + 500 OpenRice (Cantonese) to train dataset to train the model 3 and model 4 respectively. Both use the fastText embedding and positional embedding, also we change the Optimizer to Adam with a learning rate of 0.001 and dropout rate of 0.2, and the batch size is 32. For Model 5, we use the ChatGPT 3.5 turbo model to evaluate the its performance in handle the ABSA Cantonese analysis, as the model is provided by Poe, so there is no need to set the hyper-parameters. All these models evaluates its performance by the 1,000 OpenRice Evaluate dataset. The details of the models' setting are listed in Table 7.

	Model 1: Simplified Chinese Model (train by 45,000 AIC dataset)	Model 2: Simplified Chinese Model (train by 500 AIC + 500 OpenRice dataset)	Model 3: Fine-tune Cantonese Model (train by 500 AIC + 500 OpenRice dataset)	Model 4: Fine-tune Cantonese Model (train by 45,000 AIC + 500 OpenRice dataset)	Model 5: ChatGPT 3.5 turbo
Embedding	Word2vec and Positional Embedding	Both Word2vec and Positional Embedding	Both FastText and Positional Embedding	Both FastText and Positional Embedding	/
Attention stacks	3	3	3	3	/
Dimensional size	300	300	300	300	/
Hidden dimensional size	60	60	60	60	/
Optimizer	Momentum 0.7	Momentum 0.7	Adam	Adam	/
Learning rate	0.01	0.01	0.001	0.001	/
Learning rate decay	0.0	0.0	0.0	0.0	/
Batch size	32	32	32	32	/
Regularization parameter	0.001	0.001	0.001	0.001	/
Dropout rate	0.1	0.1	0.2	0.2	/

Table 7: Models' Hyper-parameters setting

To evaluate the ChatGPT 3.5 turbo performance in analyze Cantonese restaurant review, we need to write a ChatGPT prompt to predict the result of **1,000 OpenRice Evaluate dataset (Cantonese)**, in figure 17, figure 18, and figure 19, we followed the ChatGPT prompt formula to create our own prompt to obtaining the aspects' polarities from the given Cantonese restaurant review within the **1,000 OpenRice Evaluate dataset (Cantonese)**, the prompt formula usually contained the persona, context, task, exemplar, format, and tone, but only contained the Context, Task, and Instruction in our prompt:

- **Persona**: Given the scenario to ChatGPT model for obtaining a specific result, in this case we have given the scenario that is the Data Scientist as shown in figure 14.
- **Task**: Given the clear task to ChatGPT to focused on the specified task. In this case, we assigned the task to analyze the given Cantonese restaurant review to the provided each aspects' polarity as positive, negative, neutral, and not mentioned as shown in figure 14.
- **Exemplar**: Given some example to ChatGPT model to ensure the result more accurately, and the exemplar is shown in figure 15.

```
我希望你扮演資料科學家，你的任務是為廣東話評論進行情緒分析，分析評論內的每個類別的情感。["餐廳的交通是否便利", "餐廳的位置距離商圈遠近", "餐廳的位置是否容易尋找", "餐廳的服務排隊等候時間", "餐廳的服務人員態度", "餐廳的是否容易停車", "餐廳的點菜或上菜或送餐速度", "餐廳的價格水平", "餐廳的性價比", "餐廳的折扣力度", "餐廳的環境裝修狀況", "餐廳的吵雜狀況", "餐廳的用餐空間", "餐廳的環境衛生狀況", "餐廳的食物分量", "餐廳的食物味道", "餐廳的食物外觀", "餐廳的食物推薦程度", "本次消費感受", "再次消費的意願"]為留言的分析類別，情感從["正面", "中性", "負面"]裡選擇。如果不存在，回答："沒有提及"。
```

Figure 17: The ChatGPT prompt that given the Task and Persona.

```
以下是餐廳評論情緒分析的範例1：
"味道不錯的麵館，性價比也相當之高，分量很足～女生吃小份，胃口小的，可能吃不完呢。環境在麵館來說算是好的，至少看上去堂子很亮，也比較乾淨 ，一般蒼蠅館子還是比不上這個衛生狀況的。中午飯點的時候，人很多，人行道上也是要坐滿的，隔壁的冒菜館子，據說是一家，有時候也會開放出來坐吃麵的 人。"

餐廳的交通是否便利 : 沒有提及,
餐廳的位置距離商圈遠近 : 沒有提及,
餐廳的位置是否容易尋找 : 沒有提及,
餐廳的服務排隊等候時間 : 沒有提及,
餐廳的服務人員態度 : 沒有提及,
餐廳的是否容易停車 : 沒有提及,
餐廳的點菜或上菜或送餐速度 : 沒有提及,
餐廳的價格水平 : 沒有提及,
餐廳的性價比 :正面
餐廳的折扣力度 : 沒有提及,
餐廳的環境裝修狀況 :正面
餐廳的吵雜狀況 : 沒有提及,
餐廳的用餐空間 : 沒有提及,
餐廳的環境衛生狀況 :正面
餐廳的食物分量 :正面
餐廳的食物味道 :正面
餐廳的食物外觀 : 沒有提及,
餐廳的食物推薦程度 : 沒有提及,
本次消費感受 :正面
再次消費的意願 : 沒有提及
```

```
以下是餐廳評論情緒分析的範例2：
"中午飯點人比較多需要排隊～2點左右去吃的剛好可以用圍購～也不用排隊～就是平安夜店員說的都不一樣有的讓用圍購有的說不能，有點混亂～店裡的 日本店員雖然不怎麼懂中文但很熱情！確實一進店一股臭臭的味道～但是待久了就習慣了 店裡桌子前有很多圍文教怎麼使用桌上的醬料，很詳細很有愛～可以自己調自己的口兒！兩個人圍了兩個套餐～還有一份炸雞！ 吃的很飽～濃厚豚骨拉麵，加了糖心蛋、海苔和豆芽（當時沒有筍了比較遺憾）～湯頭非常濃鬱、就是單喝還是有點臭臭的，加了店裡的拉麵醬、黑胡椒和芝麻～好 很多！ 裡邊的糖心蛋特別讚！博多黑拉麵，非常推薦這個！ 沒有異味兒！ 也沒有很濃的蒜味，恰到好處非常好吃～炸雞塊剛炸出來的特別酥嫩！ 不用沾任何東西就很有味兒～就是量有點小～團購很值！以後還會來的！ "

餐廳的交通是否便利 : 沒有提及,
餐廳的位置距離商圈遠近 : 沒有提及,
餐廳的位置是否容易尋找 : 沒有提及,
餐廳的服務排隊等候時間 : 中性,
餐廳的服務人員態度 : 正面,
餐廳的是否容易停車 : 沒有提及,
餐廳的點菜或上菜或送餐速度 : 沒有提及,
餐廳的價格水平 : 沒有提及,
餐廳的性價比 : 正面,
餐廳的折扣力度 : 中性,
餐廳的環境裝修狀況 : 沒有提及,
餐廳的吵雜狀況 : 沒有提及,
餐廳的用餐空間 : 沒有提及,
餐廳的環境衛生狀況 : 負面,
餐廳的食物分量 : 負面,
餐廳的食物味道 : 中性,
餐廳的食物外觀 : 沒有提及,
餐廳的食物推薦程度 : 正面,
本次消費感受 : 正面,
再次消費的意願 : 正面
```

Figure 18: the second part of ChatGPT prompt in figure 17 that given the Exemplar

請分析以下的留言並輸出所有類別的分析結果：
試咗四樣野，蛋包飯都係得個樣，味道好一般，有偏鹹嘅情況。 要黎打卡可以 紅酒燉牛肋條 直程係難食嘅級數，完全冇牛肉味，而且一陣怪味。 完全冇胃口食落去。 綜合炸物拼盤係全部嘢比較正常嘅，一般日本餐廳級數。 結論如果想去打卡都還可以，但係唔好對食物有期待。 唔明白點解啲人吹奏話呢間嘢好食？係咪未去過日本？ 定係味蕾有問題？

Figure 19: the last part of the ChatGPT prompt in figure 17 that the target review data that to be analyzed

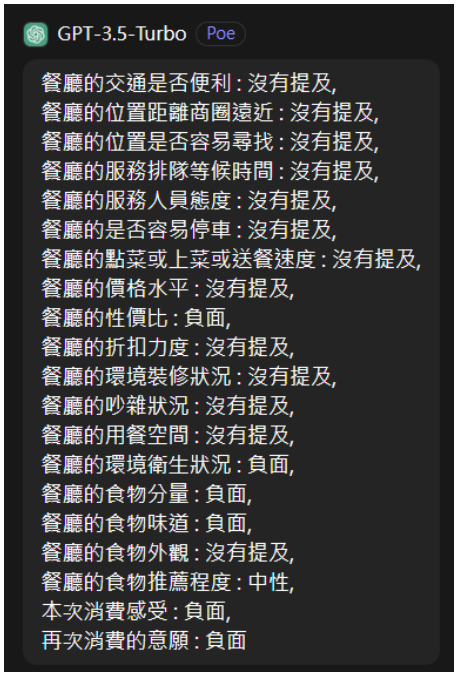


Figure 20: It shown result of the labelled data

Based on the prompt in Figure 17, 18, and 19, we can obtain the aspects and its polarities (labelled data) from the Cantonese restaurant review, and the classification result of the ChatGPT 3.5 turbo model shown in Figure 20.

id	content	location_tra	location_di	location_ea	service_wa	service_wa	service_pa	service_ser	price_level	price_cost	price_disc	environmer	environmer	environmer	environmer	dish_portio	dish_taste	dish_look	di
0	"呢個商場食店之選	-2	1	1	1	1	1	1	1	1	1	1	-2	1	1	1	1	1	
1	"將軍澳新都城日式	-2	1	1	1	1	1	1	1	1	1	1	-2	1	1	1	1	1	
2	"喺新都城中心一樓	-2	1	1	1	1	1	1	1	1	1	1	-2	1	1	1	1	1	
3	"朋友見喺將軍澳新	-2	1	1	1	1	1	1	1	1	1	1	-2	1	1	1	1	1	
4	"我平時好鐘意去呢	-2	1	1	1	1	1	1	-1	1	-2	1	-2	1	1	1	-1	1	
5	"好似係新開張咗一	-2	1	1	1	1	1	1	-1	1	-2	1	-2	1	1	1	-1	1	
6	"落單係用QR code	-2	1	1	1	1	1	1	-1	1	-2	1	-2	1	1	1	-1	1	
7	"明明已叫一小時上	-2	1	1	-1	-1	1	-1	1	1	-2	1	-2	1	1	-2	-2	-2	
8	"未開業已經好期待	1	1	1	1	1	1	1	1	1	-2	1	-2	1	1	1	1	1	
9	"睇住佢哋裝修左好	1	1	1	-1	1	1	1	1	1	-2	1	-2	1	1	1	1	1	
10	"今日咁巖女女有興	1	1	1	1	1	1	1	1	1	-2	1	-2	1	1	1	1	1	
11	"約了將軍澳的朋友	1	1	1	1	1	1	1	1	1	-2	1	-2	1	1	1	1	1	
12	"新都城好多野食	-2	1	1	-2	1	-2	1	1	1	-2	1	-2	1	1	1	1	1	

Figure 21: This figure is screen captured on the CSV dataset that stores the ChatGPT 3.5 turbo predicted result; it is the partial data in the dataset.

After that, we copy the review data and its result of Figure 20 to the dataset in the shown on Figure 21. The evaluation will be uses the correct OpenRice evaluation dataset which labelled by ours and the predicted data to evaluate the ChatGPT 3.5 turbo model performance whether suitable for handling the ABSA task in Cantonese reviews.

Models	OpenRice evaluation dataset (Cantonese or Simplified Chinese review data)	
	Accuracy	**Macro-F1**
Model 1: Simplified Chinese Model (train by 45,000 AIC dataset)	77.74 (Simplified Chinese)	40.47 (Simplified Chinese)
Model 2: Simplified Chinese Model (train by 500 AIC + 500 OpenRice dataset)	77.3 (Simplified Chinese)	31.96 (Simplified Chinese)
Model 3: Fine-tune Cantonese Model (train by 500 AIC + 500 OpenRice dataset)	76.3 (Cantonese)	30.66 (Cantonese)
Model 4: Fine-tune Cantonese Model (train by 45,000 AIC + 500 OpenRice dataset)	**78.67 (Cantonese)**	**44.76 (Cantonese)**
Model 5: ChatGPT 3.5 turbo Model	36.16 (Cantonese)	27.52 (Cantonese)

Table 8: The experimental results comparison on four different models with our OpenRice evaluation dataset.

5.6. Performance Evaluation

After the experiment, we compared 5 Models by the OpenRice evaluation dataset, as the result shown in Table 8. The model 1 has the second-highest performance in accuracy and macro-F1 scores with 77.74% and 40.47%. The model 2 has the third-highest performance in accuracy and macro-F1 scores with 77.3% and 31.96%. The model 3 has the fourth-highest performance in accuracy and macro-F1 scores with 76.3% and 30.66%. For model 4, it is the best model to handle the Cantonese restaurant review that its accuracy, and macro-F1 score with 78.67% and 44.76%. The model 5 has the lowest performance in handle the ABSA task in Cantonese restaurant reviews, its accuracy, and macro-F1 score with 36.16% and 27.52%. As a result, the Model 4 has the highest performance with other models. Therefore, we decided to deploy this model to our application to handle the problem of Aspect-Based Sentiment Analysis in Cantonese.

The lowest-performance reason for model 5 is that the ChatGPT had limited emotional intelligence and limited domain-specific expertise. ChatGPT is a language model and lacks the depth of knowledge or expertise found in domain-specific (Ray, 2023), so it cannot fully handle the responses that aspect-based sentiment analysis (ABSA).

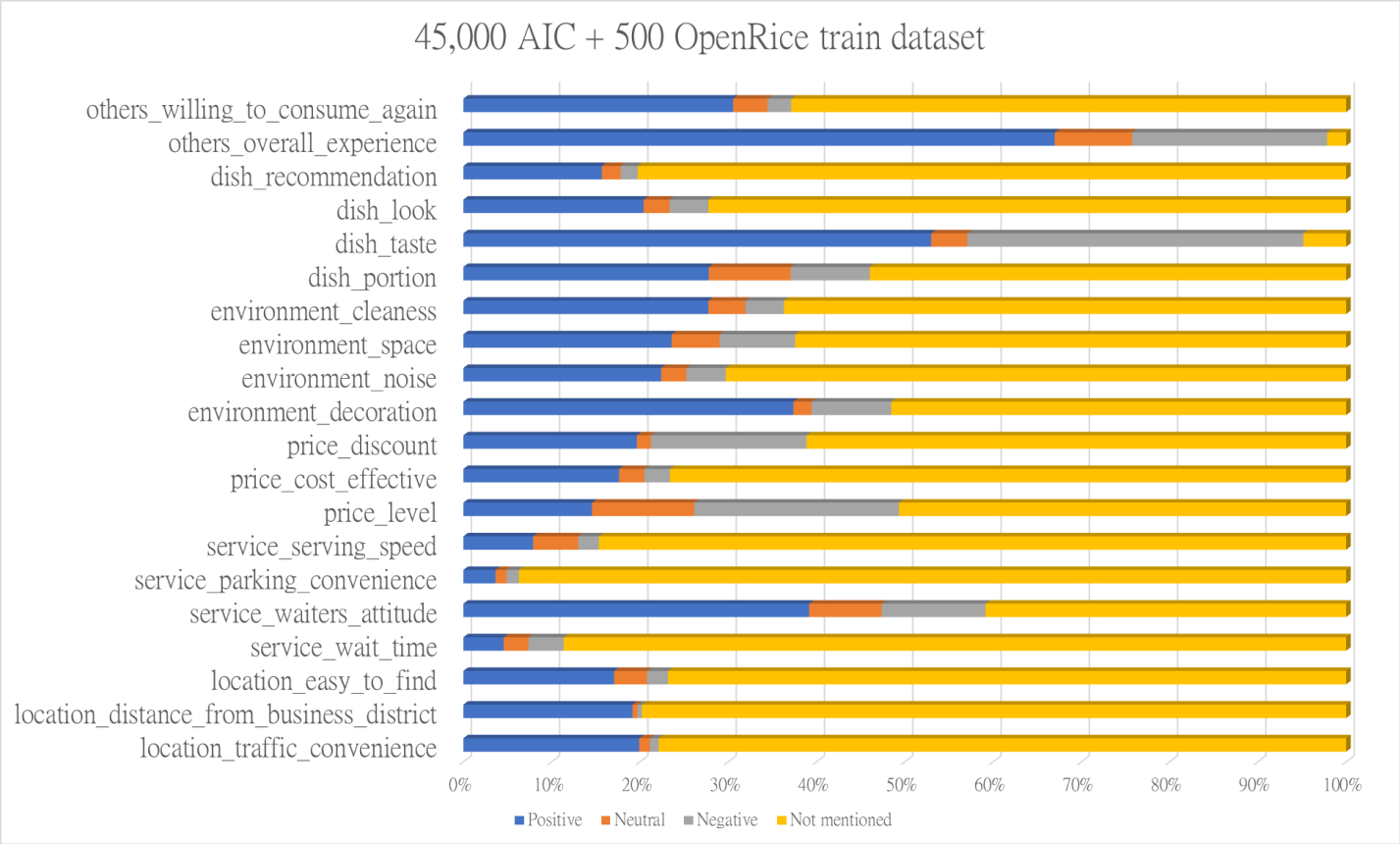


Figure 22: Data distribution for the 45,000 AIC + 500 OpenRice dataset, the blue bar represent the positive label data, the orange bar represent the neutral label data, the grey bar represent the negative data, and the yellow bar represent the not-mentioned data.

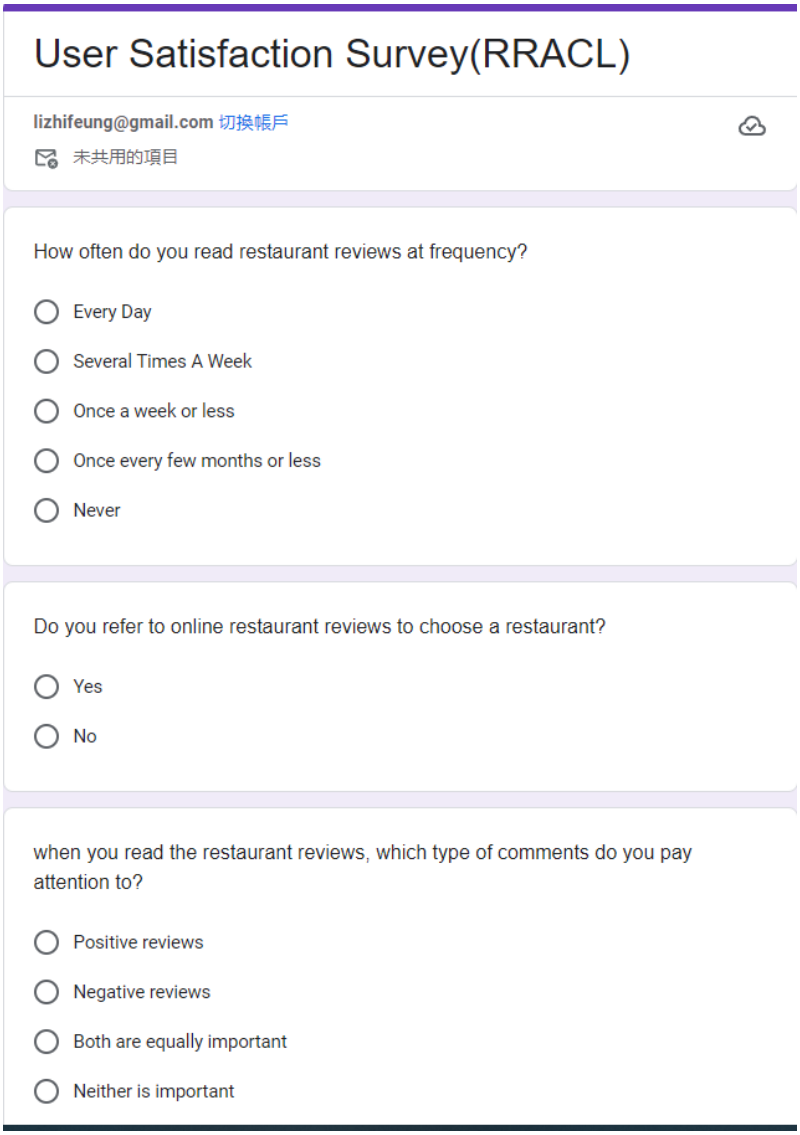
Although model 4 has the highest performance in our experiment result, its performance is not ideal since the Macro-F1 score is lower than 50%, so it still needs to improve its performance in Cantonese reviews. According to our investigation, we found that some of the labels within the "45,000 AIC + 500 OpenRice dataset" are less, in Figure 22, most of the aspects' provided very few neutral and negative labels compared with positive labels. So this will significantly affect the multiplicity of the dataset, and that will affect the model performance.

Figure 23: This screenshot is captured from the Microsoft Bing Translator, the translation result from Simplified Chinese to Cantonese. The green circle indicates the original text, and red circle indicates the wrong translation result.

Moreover, we used the Microsoft Bing Translator to translate the training and evaluation dataset, but we found that some of the translated data may make a mistake. For example, in Figure 23, most of the text translated to Cantonese is correct, the green circle on the left side is the original text which means “we order the frog", but it translated to the wrong results which means “We are frog”. That also affects the model's performance.

5.7. User Satisfaction Survey

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 24 to 26 is our survey design. Figure 27 to 36 is the survey's result; 15 people have tried our web application.



Is the web interface user-friendly?

- ◯ Convenient
- ◯ Normal
- ◯ Inconvenient

How accurate with the analysis results?

- ◯ Accurate
- ◯ Normal
- ◯ Inaccurate

How long do you think the process of analyzing the restaurant takes?

- ◯ Fast
- ◯ Normal
- ◯ Slow

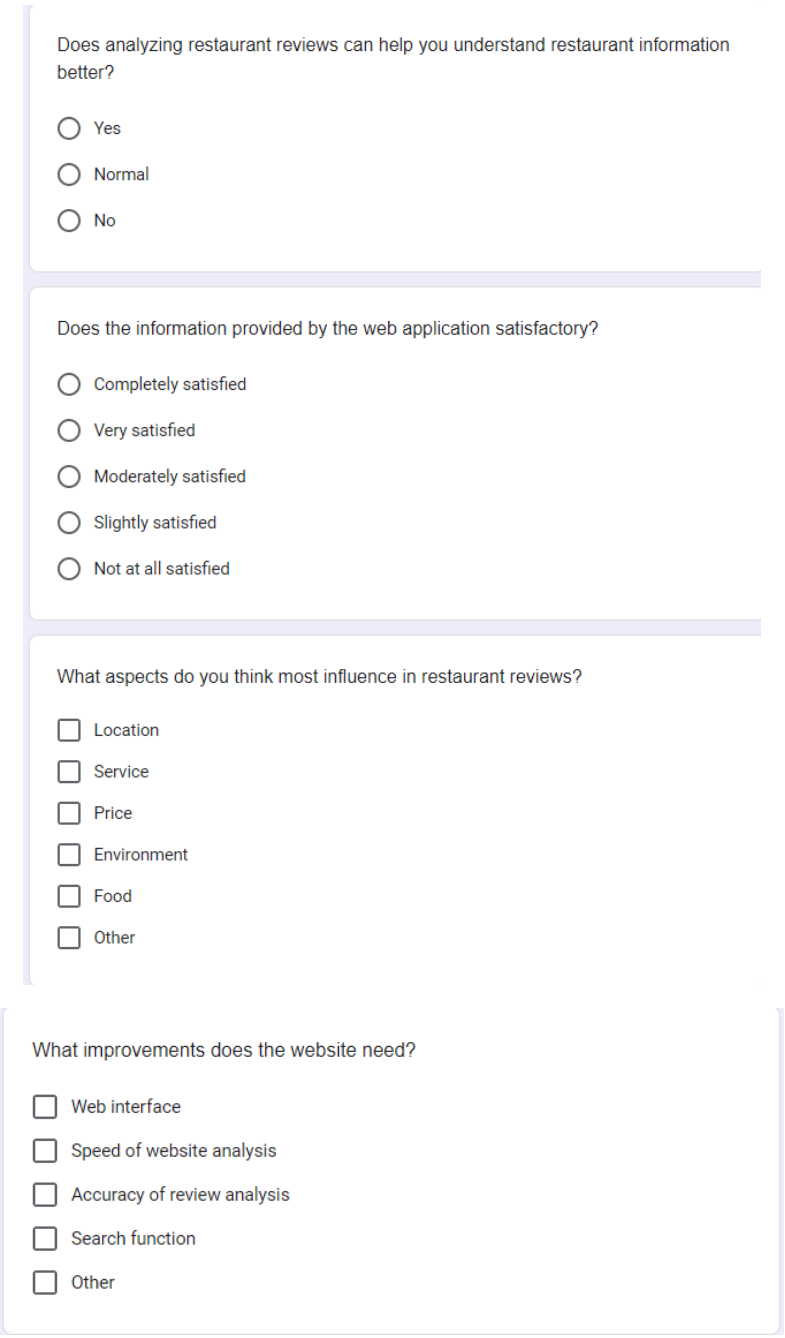


Figure 24,25,26: This is our survey to get feedback from users by using google.

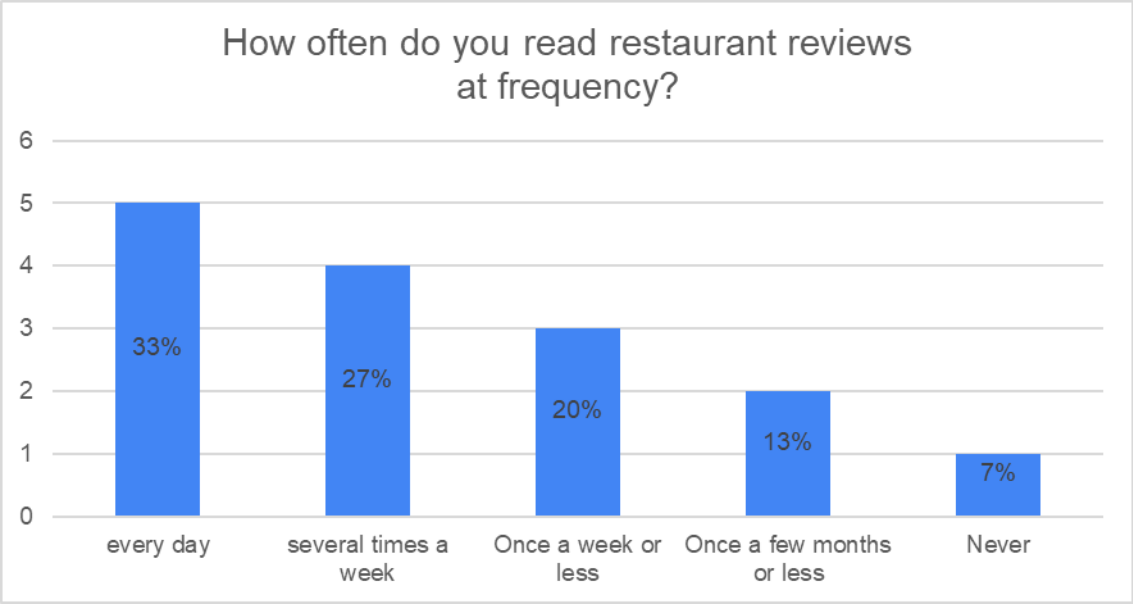


Figure 27: result of How often do you read restaurant reviews at frequency?

In this question "How often do you read restaurants reviews online", there are 5 (33%) peoples who read reviews every day, 4 (27%) peoples several times a week, 3 (20%) peoples once a week or less, 2 (13%) peoples once a few months or less, and 1 (7%) people never.

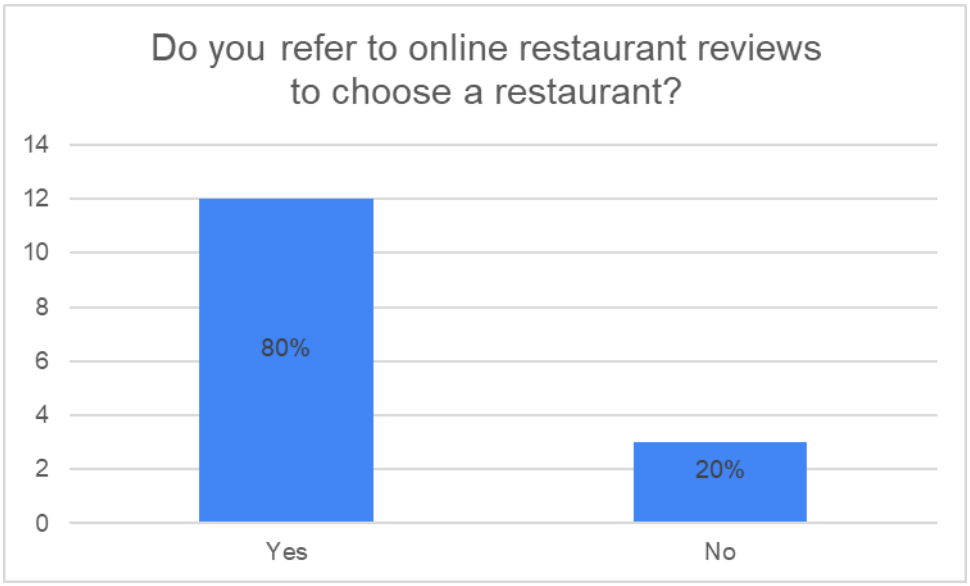


Figure 28: result of do you refer to online restaurant reviews to choose a restaurant?

In this question "do you refer to online restaurant reviews to choose a restaurant?", 12 (80%) peoples will refer to online restaurant reviews, and 2 (20%) peoples will not.

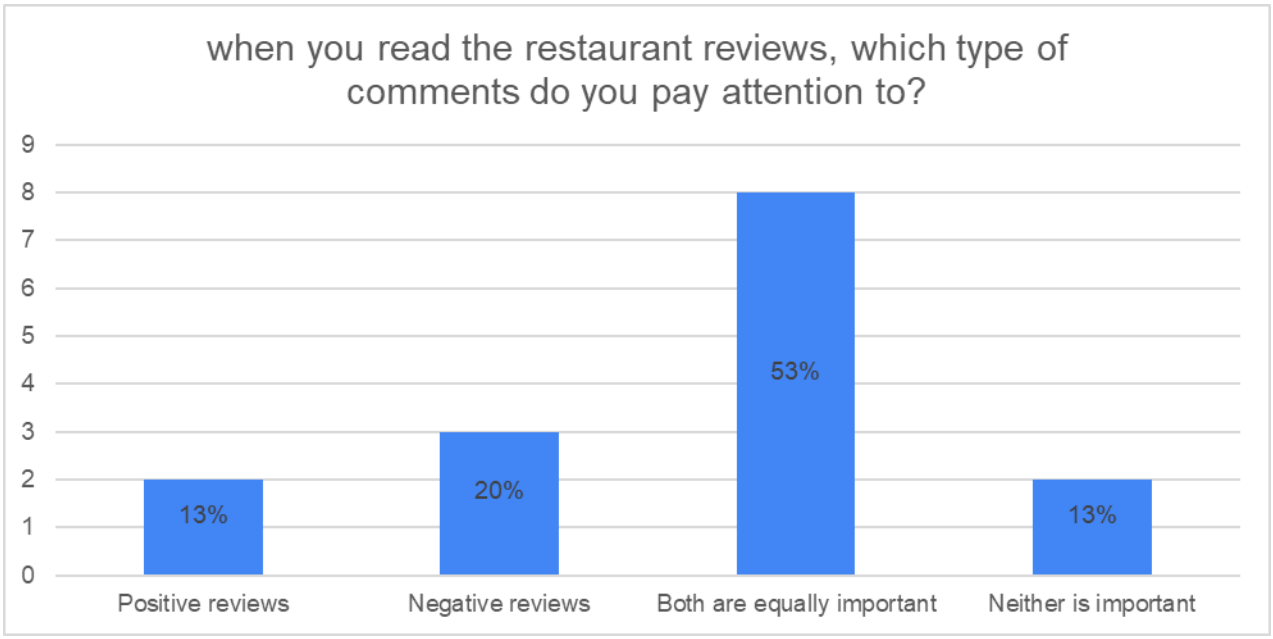


Figure 29: result of When reading restaurant reviews, which one do you pay attention to?

In the question "When reading restaurant reviews, which one do you pay attention to?", there are 8 (53%) peoples who think that both reviews are important to get attention such as positive and negative. 2 (13%) peoples think positive reviews are more attended, 3(20%) peoples think negative reviews more attend, and 2 (13%) peoples neither is important to attend.

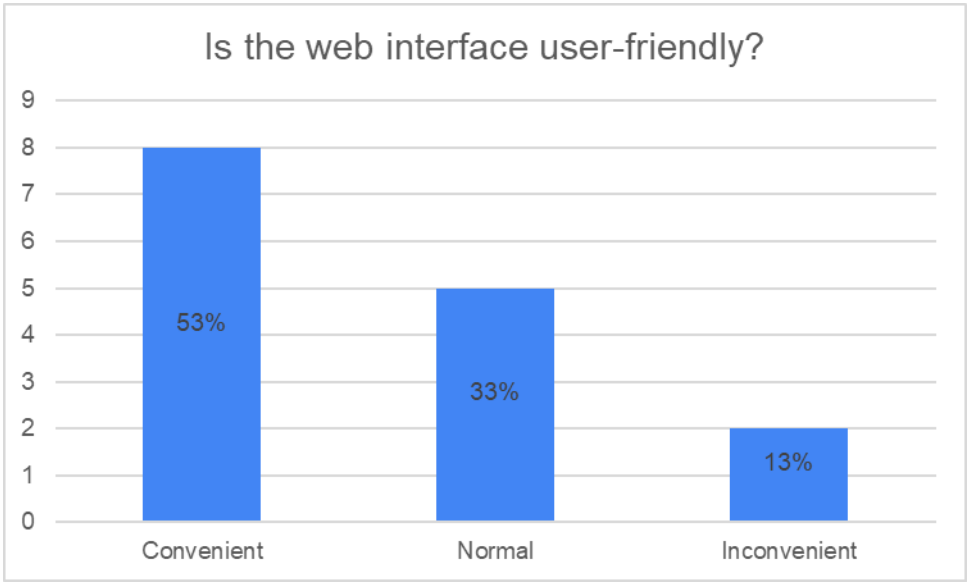


Figure 30: result of Is the web interface user-friendly?

In "Is the web interface user-friendly?" question, there are 8(53%) people who think the web interface is convenient, 5(33%) people think normal, and 2(13%) people think inconvenient.

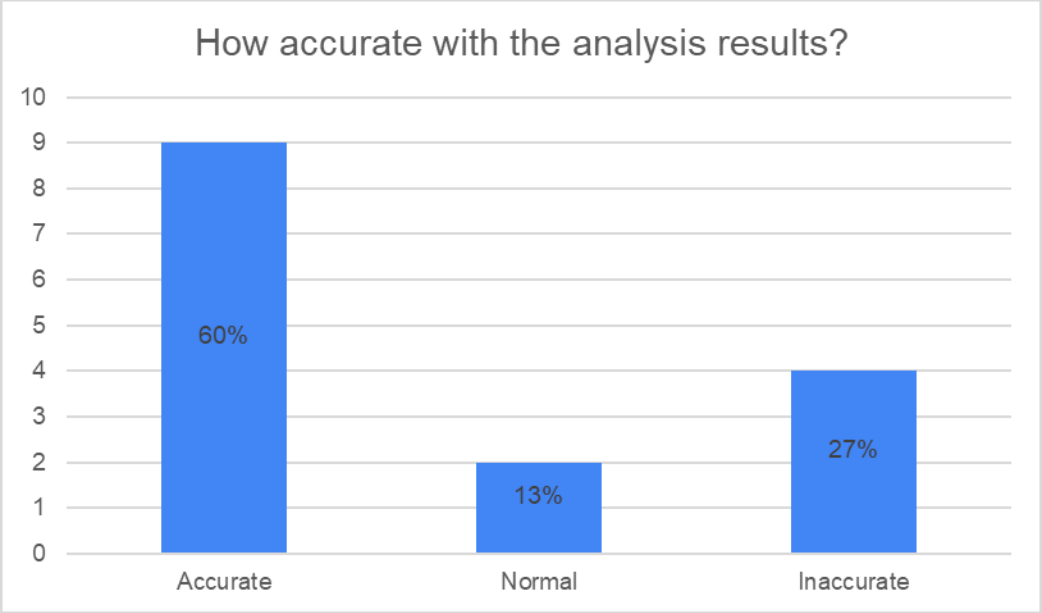


Figure 31: result of How accurate with the analysis results?

In “How accurate with the analysis results?” question, 9(60%) people think the analysis is accurate, 2(13%) people think normal, and 4(27%) people think inaccurate.

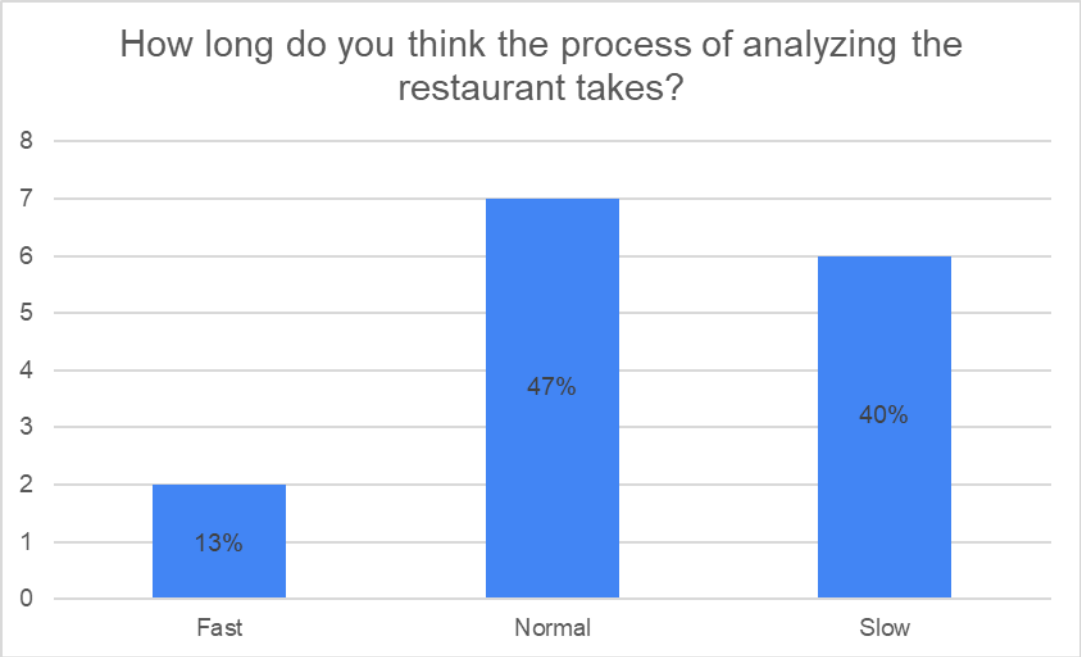


Figure 32: result of How long do you think the process of analyzing the restaurant takes?

In this question “How long do you think the process of analyzing the restaurant takes?”, 7(47%) people think normal about the analysis, 6(40%) people think is slow, and 2(13%) people think fast.

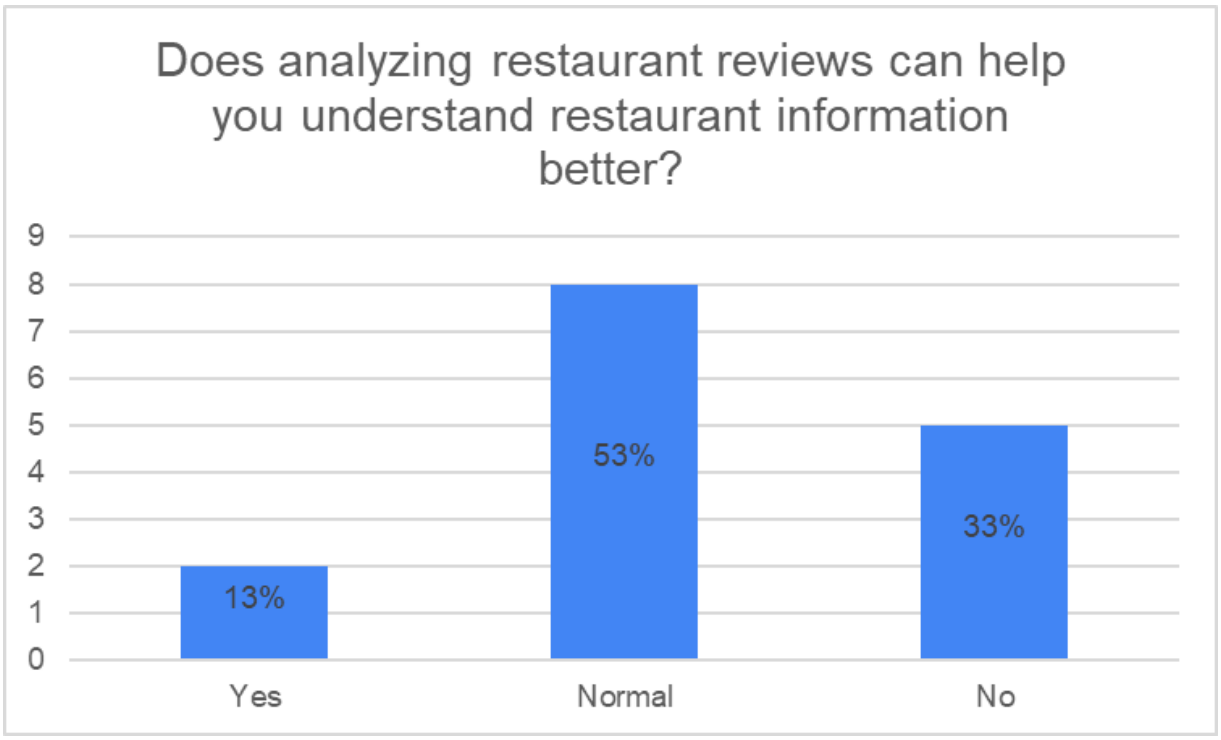


Figure 33: result of Does analyzing restaurant reviews can help you understand restaurant information better?

For the "Does analyzing restaurant reviews can help you understand restaurant information better?" question, 8(53%) people think Somewhat is helpful understand the restaurant information, 5(33%) people think can't help, and 2(13%) people think can help them to understand the restaurant information.

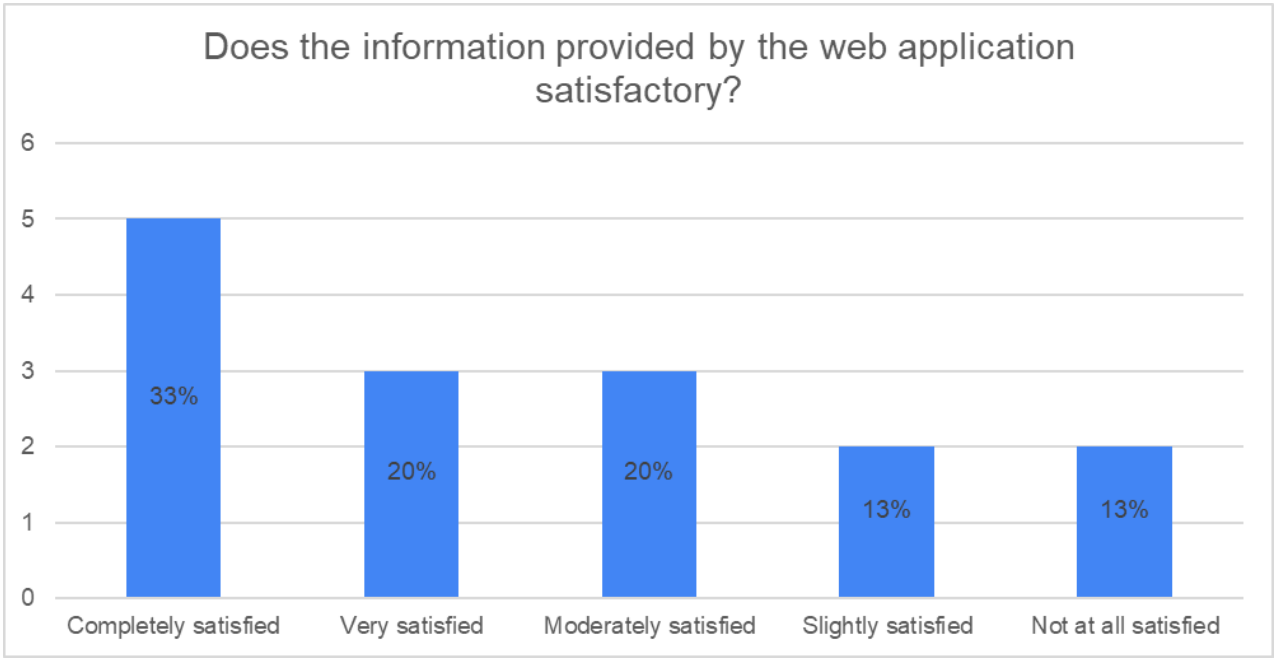


Figure 34: result of Does the information provided by the web application satisfactory?

In this question "Does the information provided by the web application satisfactory?", there are 5(33%) people feel completely satisfied for our web application, 3(20%) people feel very satisfied, 3(20%) people feel moderately satisfied, 2(13%) people feel slightly satisfied, and 2(13%) people feel not at all satisfied.

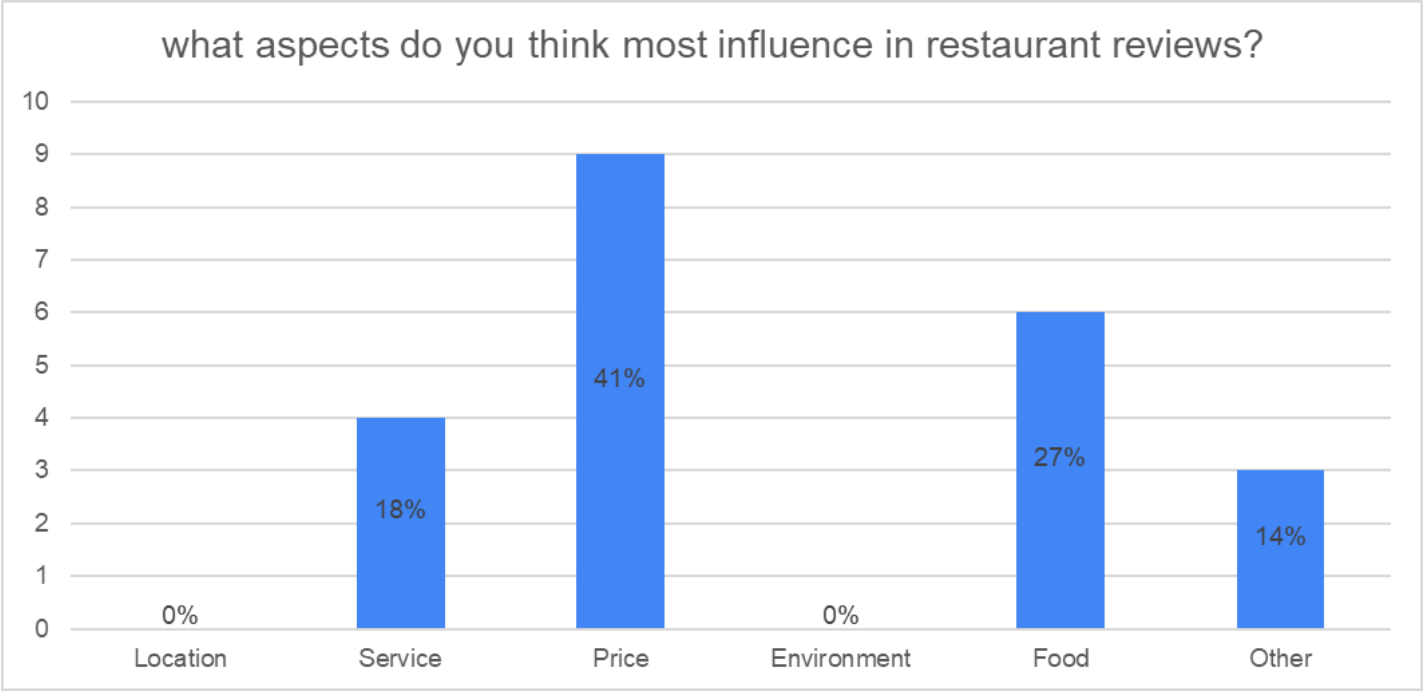


Figure 35: result of what aspects do you think most influence in restaurant reviews?

For the question "what aspects do you think most influence in restaurant reviews?", 9(41%) people think more influence is price, 4(18%) people think service is influence, 6(27%) people think food is influence, and 3(14%) is think other.

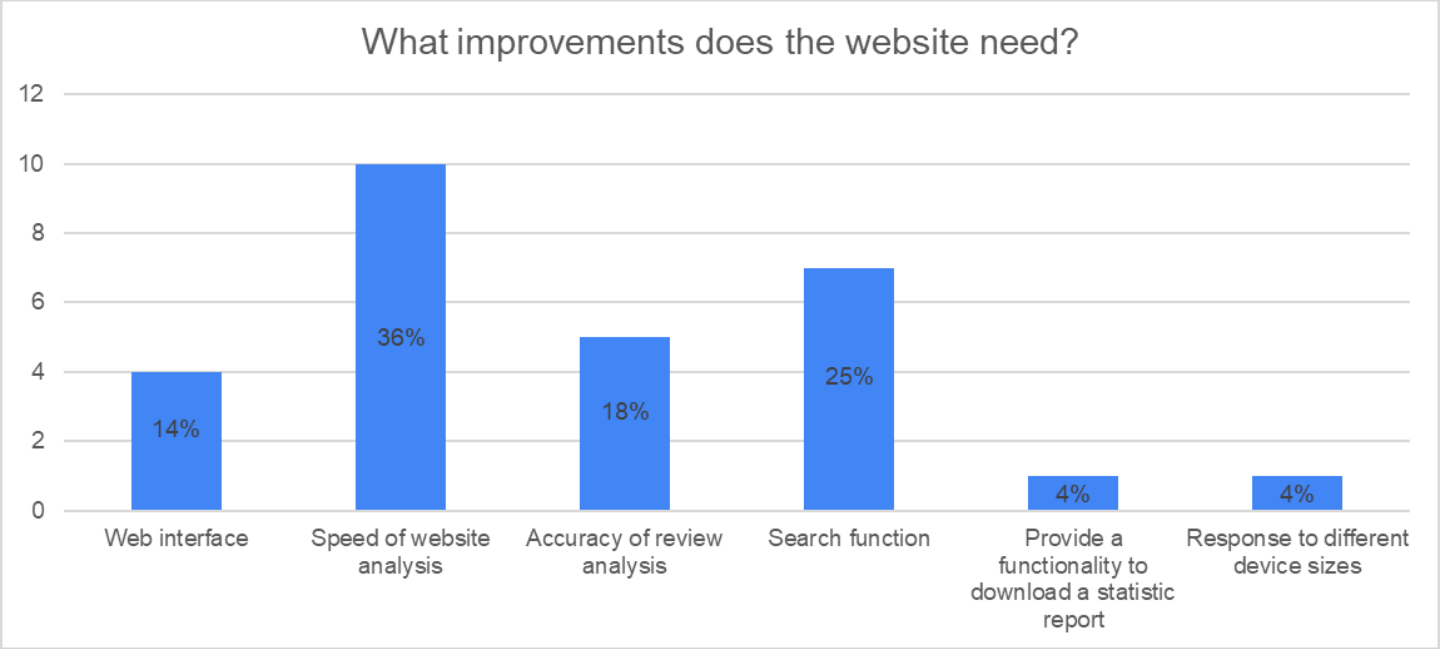


Figure 36: result of What improvements does the website need?

In this question "What improvements does the website need?", 10(36%) people think speed of website analysis need to improve, 7(25%) people think search function needs to improve, 4(14%) people think web interface needs to improve, 5(18%) people think accuracy of review analysis needs to improve, 1(4%) person think need provide a functionality to download a statistic report, and 1(1%) person think need to response to different device sizes.

5.8. User Evaluation

According to the survey results, it was found that 80% of users rely on online restaurant reviews when making decisions about where to dine. Conversely, 20% of respondents stated that they do not feel the need to look up information about restaurants before dining. Among the users who do seek information beforehand, 53% of them indicated that they consider both positive and negative reviews in order to make an informed choice. This suggests that they take into account the strengths and weaknesses of a restaurant before finalizing their decision. On the other hand, 33% of participants expressed complete satisfaction with the information provided by the web application, indicating a high level of trust in the reviews and ratings available. Moreover, 60% of respondents stated that our system provided accurate analysis results.

However, there are areas in the web application that require improvement. Approximately 36% of respondents believe that the speed of the web application's review analysis needs to be enhanced. This implies that the process of retrieving and analyzing reviews may currently be slow or inefficient, potentially impacting the user experience and there are 27% of the users think the web application is inaccurate base on Figure 31. On Figure 35, very few users would look into the location and environment category, The underlying factors for this observation may be attributed to both of the aspect is inaccurate. As for the user-interface there is still room for improvement, about 14% suggested our result page has too many graphs as our results. These findings underscore the significant role that online restaurant reviews play in the decision-making process for a majority of users. To enhance the overall user experience and address the identified areas for improvement, it is recommended to prioritize enhancing the speed and efficiency of the web application's review analysis. By doing so, the application can better cater to the needs and expectations of users who heavily rely on online reviews when selecting a restaurant.

6. Conclusion

In conclusion, this report presents an application, named “Restaurant Review Analyzer for Chinese Language (RRACL)” which aims to help restaurant owners to better collect opinions from customer(s), also, provide a way for the restaurant customers to see more accurate rating of the restaurant. The RRACL provided the functionality to analyze the Cantonese restaurant reviews that are obtained from OpenRice, it predicted result will show on the graphical user interface with the statistic charts. This tool deployed the Cantonese Aspect-Based Sentiment Analysis (ABSA) model that is used to predict the sentiments for restaurant reviews in each aspect. The objectives of this report have been achieved which are the Cantonese dataset creation, model adaptation, model training and evaluation, web application development and model deployment, and collect opinions by user satisfaction survey.

To achieve these objectives, we manually collected 1,500 review data from OpenRice and labelled the sentiments for each aspect, and data was separated into a 500-training dataset and a 1,000-evaluation dataset. Also, we collected an AI Challenger Sentiment Analysis Training set 2018 dataset from the internet and only keep first 45,000 reviews. For these datasets, we have created 4 sets of data to train and evaluate and through "Microsoft Bing Translator" to translate the datasets to either Simplified Chinese or Cantonese. Moreover, we use the Original AI Challenger ABSA model to setup 2 experiment environments, and we also adapted this model for Simplified Chinese to handle the Cantonese dataset. The adapted model will also set up 2 experiment environments. Addition, we used the ChatGPT 3.5 turbo model as the fifth experiment environment. These environments was use the 1,000-evaluation dataset to evaluate the performance, as the result, we decided to deploy the Model 4 to application since this model has the highest performance. In addition, we developed the RRACL web application to deploy our evaluated model to perform the analysis task in Cantonese. Lastly, we distributed RRACL systems to 15 test users to test and fill out the survey to evaluate the effectiveness, 33% of test users think our system provided enough information about the restaurant, and 60% of the test users think the analysis results are accurate.

However, the RRACL application has some limitations. We found that 27% of the test users stated the analyzed results are inaccurate, especially the Location aspects and Service aspects, also, the macro-F1 score of the model 4 was lower than 50% the underlying factors for this observation may be attributed to several reasons. Firstly, some aspects' labels within the training dataset were imbalanced, the majority data that were labeled were positive and not mentioned within the dataset, which means the negative label and neutral label are the minority. Secondly, the Microsoft Bing Translator sometimes returned inaccurate results. These problems will cause the model's performance to be low. On the other hand, about 36% of test users stated that the analysis speed of RRACL is slow, the reason that may related to the model prediction time. Moreover, as we only deployed a Cantonese model to RRACL, which means the application focuses on handling the Cantonese reviews, analyze the Simplified Chinese may have returned the wrong results.

To solve the limitations, the future work should mainly focus on enhancing the model’s performance. The extension for the training dataset may be a good direction to enhance the model’s performance in handling Cantonese reviews. Also, the researchers will be encouraged to explore alternative approaches, to enhance the accuracy of translate the review data of the AI Challenger Sentiment Analysis Training set 2018 from Simplified Chinese to Cantonese. Further, the researchers may enhance or adapt the Cantonese model to reduce the prediction time. Moreover, it may consider supporting the multi-models within an application to handle different language reviews such as English and Simplified Chinese. In Addition, the researchers may focus on utilizing the ChatGPT prompt to handle the ABSA tasks in restaurant reviews.

7. References

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

BYVoid. (2023, October 15). *OpenCC*. Retrieved from GitHub: https://github.com/BYVoid/OpenCC

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

Lee, J., Chen, L., Lam, C., Lau, C. M., & Tsui, T.-H. (2022). PyCantonese: Cantonese Linguistics and NLP in Python. *Proceedings of the Thirteenth Language Resources and Evaluation Conference* (pp. 6607–6611). European Language Resources Association.

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.

Raganato, A., Scherrer, Y., & Tiedemann, J. (2020). Fixed Encoder Self-Attention Patterns in. *Findings of the Association for Computational Linguistics: EMNLP 2020* (pp. 556-568). Association for Computational Linguistics.

Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems, 3*, 121-154.

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

Su, J. (2023, August 1). *Master the Perfect ChatGPT Prompt Formula (in just 8 minutes)!* Retrieved from YouTube: https://www.youtube.com/watch?v=jC4v5AS4RIM

Text Translation. (n.d.). (Microsoft) Retrieved from Microsoft: https://www.microsoft.com/en-us/translator/business/translator-api/

ToastyNews, & lso. (2020, July). *hong-kong-fastText*. Retrieved from github: https://github.com/toastynews/hong-kong-fastText

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

Use Font Awesome on the Web. (n.d.). Retrieved from Font Awesome Docs: https://docs.fontawesome.com/web/

wchan757. (2019, November 30). *Cantonese Word Segmentation*. Retrieved from github: https://github.com/wchan757/Cantonese_Word_Segmentation

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 Cross-Entropy Loss Function? (2023, June 15). (The 365 Team) Retrieved from 365 DataScience: https://365datascience.com/tutorials/machine-learning-tutorials/cross-entropy-loss/#1

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/

Yegulalp, S. (2024, January 5). *What is TensorFlow? The machine learning library explained*. Retrieved from InfoWorld: https://www.infoworld.com/article/3278008/what-is-tensorflow-the-machine-learning-library-explained.html

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

8. Appendixes

Appendix A. 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) 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	15 March 2024
Datasets creation: Label the aspects for each review (Stage 2)	Yeung Ho Yin Tommy, Wong Ping Kuen, LI Chi Fung, Rai Jon	15 March 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	30 March 2024
AI: Models Evaluation (Stage 3)	Yeung Ho Yin Tommy, Wong Ping Kuen	30 March 2024
Web Application: UI design (Stage 1)	Li Chi Fung, Rai Jon, Yeung Ho Yin Tommy	10 April 2024
Web Application: UI implementation (Stage 2)	Li Chi Fung Rai Jon Yeung Ho Yin Tommy Wong Ping Kuen	10 April 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	10 April 2024
Web Application: Data Analysis (Stage 5)	Wong Ping Kuen, Yeung Ho Yin Tommy	5 Apr 2024
Survey design	LI Chi Fung Rai Jon	20 Jan 2024