COMP4801 Final Year Project
An online food catalogue based on open crowdsourcing (Group 2)
Detailed Project Plan

Group Members
LAI CHEUK HIN (3035186152)
LEUNG LOK MING (3035125031)

Supervisor
DR. CHIM T W
The University of Hong Kong
1 Background

Dining is one of the most important concerns of the citizens nowadays. There are many online food catalogues for people to search for foods and restaurants. However, most of these food catalogues relying on the cooperation between platform and the restaurants, all of the information need to be updated manually. This kind of business strategy requires lots of manpower, it may not be capable for the companies in smaller scale.

Also, especially in Hong Kong, choices of restaurants are sometimes too many, making us feel difficult to decide where to dine. Even if we use those food catalogues or reviews application, it is still inconvenience to quickly scan through the foods from different restaurants.
2 Previous Work and Related Studies

2.1 Openrice
Openrice is a famous food and restaurant guide in Hong Kong and some other Asian areas, it provides both website and mobile application for the users. Openrice’s restaurant search engine allows users to search from its database that includes over 40,000 restaurants in Hong Kong [1]. Users can also rate and write reviews on restaurants to provide first-hand information for the others. There is photo upload function for user to share the photos of food and menu. However, it does not analyze the information contained in the photos. Therefore, the food in photos can not be recognized and classified by the system automatically.

2.2 Yelp
Similar to Openrice, Yelp is a crowd-sourced review website/mobile application for restaurants around the world. With a stronger user base, Yelp has provided the dataset for developers to do their analysis [2]. And, in fact, researches about using reviews from Yelp dataset to improve the restaurant has been done [3], showing the potential of the data of reviews. By downscaling the precision of the review from restaurants to meals, the data collected should be able to reveal more information from the customers. Hence, the food catalogue in this project should have potentials for data mining as well.

Also, Yelp is recently developing an AI to recognize the user’s photo [4]. The technology is planned to use at photo beautification and auto-chopping. With the current APIs such as Google Vision, we could also build the similar function and perhaps use the technology in a different aspect.

3 Objective
This project is aimed to create an open-source platform for recording the prices and reviews of restaurants in Hong Kong. The users can know the most updated information about restaurants immediately by opening their app. A photo recognition function should also be built in the system in order to shorten the time needed for inputting the data, including the menu and the photos for dishes.

The ultimate goal of the project is to collect a copious amount of data related to dishes in Hong Kong and use them as the dataset for data mining as currently there is no dataset related to each single dishes in Hong Kong. There is the potential for analysing food and restaurant preference of Hong Kong people.

4 Scope
In this project, an iOS application for viewing and editing the food catalogue will be worked on. For simplicity, only restaurants in Hong Kong will be recorded
in the system. Also, the system will be focused on the restaurants with normal order system, but not including those in the form of buffet, for the sake of simplicity in database design. The review system will only be in the form of scoring. Furthermore, depending on the performance of the actual implementation of the photo recognition function, suggestions of the food name may be provided by the system.

5 Methodology

5.1 Function List

Membership system

Users can register as a member of the application. By logging into the system, the user the store the record of dining including the photos. Also, registered users can review their meals.
Restaurant Searching

Users can search the restaurants nearby. The application will suggest the recommended dishes for each restaurant available (See Figure 1).

Menu viewing/ editing

Users can view the menu provided and organized by other users. The information will be shown in the form of text instead of photo. Users can also upload the photo of the menu. The application will try to recognize the texts on the photo and convert them into texts.

Photo uploading

Users can upload photos of the dishes they have tasted. The application will try to recognize which dish is that by analyzing the image.

Restaurants/ Dishes recommendation

Based on the dining habit of the user, the system will suggest restaurants that fit the user.

5.2 Front-end application

Swift 3 is chosen as the language for developing the frontend iOS application in this project. Comparing to Objective C, Swift have higher performance in terms of the program speed according to Apple’s research [5]. Also, Swift has a safe programming pattern due to the strict rules of optional types. This pattern makes clear nil optional values. While Objective C can call a method containing a nil pointer variable, this may lead to the app working improperly in some situations. Therefore, the time needed for bug fixing in Swift would be less when comparing to Objective C.

For the design pattern of this application, Model-View-Presenter (MVP) will be applied. It is a variant from the typical pattern Model-View-Controller (MVC). In the traditional Cocoa MVC design pattern suggested for iOS applications, it is hard to separate the View and Controller because the View Controllers are closely involved in View’s life cycle, it is difficult to mock the Views in unit tests. Therefore, the testability of MVC is bad due to the tightly coupled View and Controller. On the other hand, in the MVP pattern, there is a Presenter that contains the UI business logic for the View, and the View here is generally passive (see Figure 2). There is no layout code in the Presenter and it is only responsible for updating the View with data and state. As the interaction between Presenter and View is through an interface, it is easier mock the View and perform unit tests than in MVC.
5.3 Database system and server

As the project will be implemented in the form of a iOS application with a connection to the server, when the user has to upload or retrieving any data from the database, the server has to work as a medium to obtain the data from the database and then send it back to the user. Therefore, a reliable and fast database system and server is essential. The database system will be implemented by MongoDB and run on the Amazon Relational Database Service by AWS (Amazon Web Service). For the server, it will be implemented by node.js and also ran on the service provided by AWS.

5.4 Computer Vision

The technology of computer vision will be used in two aspect - photo recognition and text recognition. To implement this, Google Cloud Vision API and Microsoft Azure Computer Vision API will be considered. Both of the APIs have the image analysis function, providing categories calculated by the computer, and the optical character recognition (OCR) function, converting characters on images to text data.

For the OCR function, Google Cloud Vision API will be implemented as it included the multi-language recognition function, while Microsoft Azure can only used for English language.

For the photo recognition function, we will decide to use one of them after testing their accuracy and quality of the object recognition. The reason of using existing API is that building the foundation of deep learning computer vision requires an extraordinary computational power and huge amount of raw data. Considering the scope of this project, building the system from scratch is not applicable. Therefore, using existing APIs would be the most reasonable way.

5.5 Machine Learning

Machine Learning technique will be implemented for the recommendation function and the photo recognition function. Most of the computation will be executed in the server side for the ease of accessing the data. To actually implementing the machine learning technique into the application, library like Limdul.js will be considered as the main structure of the component. By constantly storing and retrieving the input and output result of the algorithm, the system should become more accurate on identifying dishes.
5.6 UI Design

The homepage is mainly consist of a navigation bar, a search bar, a list of food recommendations and a tab bar (see Figure 3). The app will list out photo and name of the foods based on the user’s location and ranking in the system. In order to keep the app simple and easy to use, all of the main functions can be accessed directly from the home page. For instance, users can tap the button on the navigation bar to view nearby restaurants in a mapview, search restaurants by typing in keywords in the search bar, take photos write reviews by tapping the camera button in the tab bar at the bottom.

An interactive mockup of the application can be viewed on the project website or at the link below:

https://xd.adobe.com/view/c1964720-34c4-47ca-9c00-f84665565fc6

Figure 3: Mockup homepage of the app
6 Business Potential

According to the research done by Hong Kong Trade Development Council [6], the advertising expenditure of 2015 in mobile application rose dramatically by 51% from previous year, which shows that the advertisers in Hong Kong are willing to spend more on mobile advertisement. At the meanwhile, the Adspend report from admanGo [7] stated that the beverage industry is ranked the fourth by Total Media Ad Spending in Hong Kong. The total spending is about 600 millions Hong Kong dollars. The statistic above shows that there is a large business potential for our application, as it is an application on mobile platform that related to foods and drinks.

As an open crowdsourcing application, if we can fascinate numerous users, it is not only a benefit for collecting information, but also an advantage to attract companies to place advertisements. In case a restaurant is interested to place advertisement, we could provide 2 offers for them to increase their exposure in the application. One is to boost the ranking of the restaurant such that it is easier to be found by the users. The other way is to place an advertisement directly on a pop-up page, then the users must watch it when they use the application.

7 Schedule and Milestones

<table>
<thead>
<tr>
<th>Phase</th>
<th>Deliverables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase 1 (Design)</td>
<td>• Detailed project plan</td>
</tr>
<tr>
<td></td>
<td>• Project website</td>
</tr>
<tr>
<td>Phase 2 (Implementation)</td>
<td>• Preliminary implementation</td>
</tr>
<tr>
<td></td>
<td>• First presentation</td>
</tr>
<tr>
<td></td>
<td>• Intermediate report</td>
</tr>
<tr>
<td>Phase 3 (Testing and finalize)</td>
<td>• Finalized tested implementation</td>
</tr>
<tr>
<td></td>
<td>• Final presentation</td>
</tr>
<tr>
<td></td>
<td>• Final report</td>
</tr>
<tr>
<td></td>
<td>• Project exhibition</td>
</tr>
</tbody>
</table>
8 Timeline

2017 September
• First meeting with supervisor
• Study related technologies

2017 October
• Submit Deliverables of Phase 1 (1 October 2017)
• Detailed project plan
• Project website
• Frontend UI/UX design
• Database design

2017 November
• Implement basic functions
• Setup database system and server

2017 December
• Implement object recognizing function
• Mid-term review

2018 January
• First presentation (8-12 January 2018)
• Submit Deliverables of Phase 2 (21 January 2018)
• Preliminary implementation
• Detailed interim report

February
• Implement Optical character recognition

March
• Test and Fine tune the application
April

- Submit Deliverables of Phase 3 (15 April 2018)
- Finalized tested implementation
- Final report
- Final presentation (16-20 April 2018)

May

- Project exhibition (2 May 2018)

9 Risk, Challenge and Mitigation

9.1 Implementing photo recognition function

There exists a potential risk that the computer vision API provided by Google Cloud or Microsoft Azure could not identify the object in the photo with a high accuracy. Since the categories available in both the APIs is not designed for food only, we may have to further analysis the categories data returned by the API and process it with a library of food categories made by our own. In other words, it may involve a further machine learning process in the application.

The application should be able to compare the analysis result of the API with other previous results, then provide several suggestions based on the dishes in that restaurant. To implement this, we are planning to use the machine learning library such as the Limdu.js library for javascript.

9.2 Implementing optical character recognition (OCR) function

In this application, OCR will be used as a tool for user to input the data of the menu. The main challenge of implementing this function is that the OCR API is not guaranteed to be accurate. It would easily be affected by the lighting and angle when the user takes the picture. Therefore, the system will not be fully relied on the OCR. Users should be able to edit and adjust the contents after they have taken the photo.

9.3 Inaccurate data entry and improper user behaviour

This open crowdsourcing application is going to accept data entries from a large population of participants. There will be a great opportunity for the system to receive inaccurate data entries. To mitigate this risk, a voting system will be implemented. As it will be resource consuming to validate all obtained information, the system will be trust-based, assuming all data entries are correct.
If a data entry is flawed, the users can vote down to it. As a result that entry’s ranking will be lower and there will be less chance for the other users to see it.

A reporting system will also be implement to handle a more serious situation. If a user upload any information that contain improper content, the other users can report this to the system. The system will check if the photo contain any adult content by the vision API. If the system detect improper content from the photo or there is a number of reports towards a data entry, that data entry will be suspended and hid until the verification by system administrator.

10 Conclusion

This project plan reveal the basic structure and the aim of the food catalogue application. With the help of the existing technology, integration of data and machine learning will be applied in the application. Difficulties may be encountered when we implement the algorithms, yet we will try to achieve the aim of the project - creating a convenience and easy-to-use application to look for menus.

References


