Content-based Recommendation by Computer Vision

Authors: Hui Chen, Ruoshi Zhao, Lejia Hu

Introduction

Meet fresh business background

Meet Fresh is a Taiwanese dessert restaurant chain, with locations in Asia, Australia, New Zealand, UK, Canada and the United States. Corporate headquarters are in New Taipei City. Founded in 2007, the chain specializes in fresh Taiwanese desserts, including soft taro balls and herbed jelly. Now Meet Fresh wants to expand its North America market, and already has 42 restaurants opening in various locations in the United States. However, since Meet Fresh is focusing on bringing authentic Taiwanese street food to the US market, and does not know about how people live in the United States like their food, they need to give these first time users a recommendation when they are ordering.

Project background

As a result, our project aims to give our customers a menu recommendation on web api. Only by answering a few questions, the users would be able to get recommendations from our website.

Our last project already covered a pre-ordering recommendation based on collaborative filtering, in this project we aim to improve it by adding a post-ordering recommendation based on content. In this way we can make the system more complete and also enhance the purchase.

Market Analysis

In our market research, it shows that many apps, for example, uber eats would give both pre-ordering recommendation (most people like, or chef recommend), they also provide post-ordering, for example, before going to check out, uber eats would suggest users to take something extra, such as drinks or other foods. In this way, users may take something extra and promote the sales of the app. We would like to do the same for MeetFresh as well, to increase the sales as well as user experience.

Final product solution design before the technical development

The product solution is divided into two parts, a recommendation engine, and an enhanced recommendation engine.

For the recommendation engine, we would provide a one-stop shop, easy, and unbiased recommendation system for the user without obtaining so much information from the user as the final product solution before technical development. The frontend would be some responsive app where the user can run our recommendation across different platform or device. It was the collaborative filtering technique that we were looking for in the first place in this case. It is essential to have the user's MeetFresh purchase history as the baseline for the collaborative filtering technique. In this way, we can map a new user's feature matrix to an existing user's feature matrix based on similarity. Therefore, the user's purchase history is our target data source directory. Furthermore, we would like to corporate with Natural Language Processing (NLP) feature into our recommendation engine so that users would be aware of the product rating of the item. In this case, *Yelp* is our primary data source directory. However, one thing needs to keep in mind is the reviews on *Yelp* are biased since the people submitting them are not randomly selected; rather, they themselves have taken the initiative to write. This leads to self-selection bias the people motivated to write reviews may have had poor experiences, may have an association with the establishment, or may simply be a different type of person from those who do not write reviews. Therefore, further investigation needs to be done about NLP feature.

Following the addition of an item to the cart, the enhanced recommendation engine further recommends other products. Visual representations of the products would replace text in this section to enhance the user's experience. Before the users check out their cart, we recommend similar products based on their existing cart item(s). Images of products on the menu (both website and hard-copy menu) would be the ideal data source without any copyright issues.

Product technical solution design

- <u>What data to collect?</u>
  For the data collection, we utilized the existing user's MeetFresh purchase history during the survey through the NeedFinding step. During the survey, 92 participants participated the the survey and completed our purchase history data source.

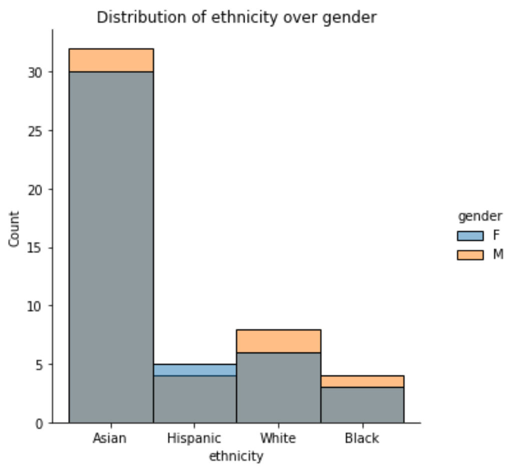

Fig 1: Distribution of ethnicity over gender on survey dataset

As we can see from Fig 1, the dataset is heavily sampled with “Asian” ethnicity. At the same time, male takes over large sampling than the female in the dataset.

Furthermore, we managed to get the MeetFresh menu data (food and drink) and below is the class distribution of the food and drink menu in different categories:

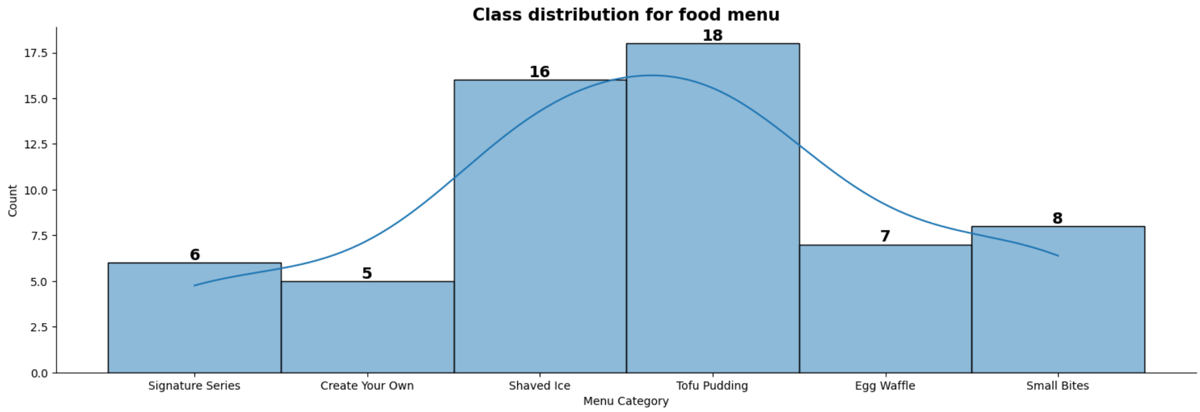

Fig 2: class distribution of Food Menu

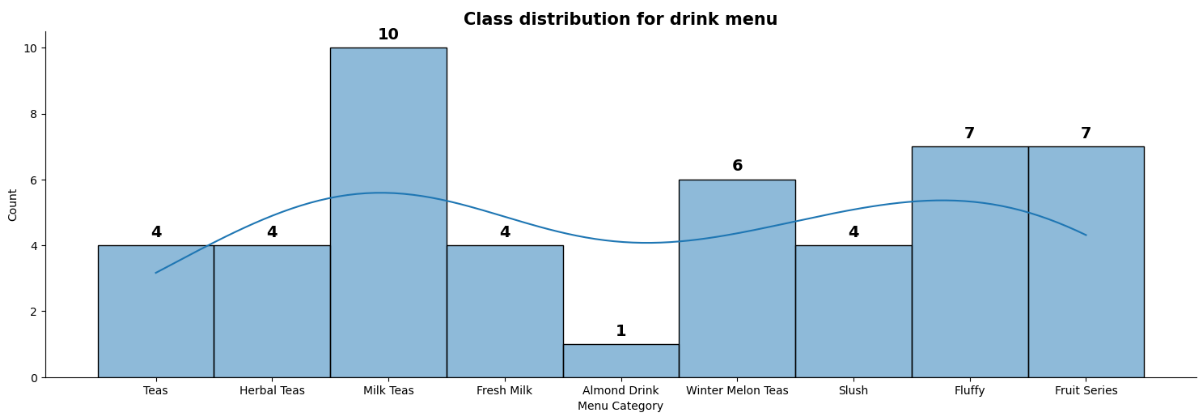

Fig 3: class distribution of Drink Menu

Lastly, we utilized web-scraping technique to scrape product menu image data as our image data primary data source directory. And below is the sample of what we have collected:

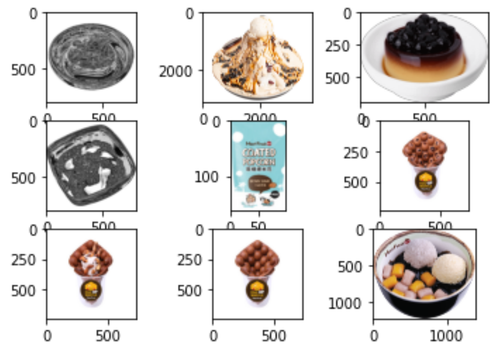

Fig 4: sample image of MeetFresh product

And the class distribution of image is shown below:

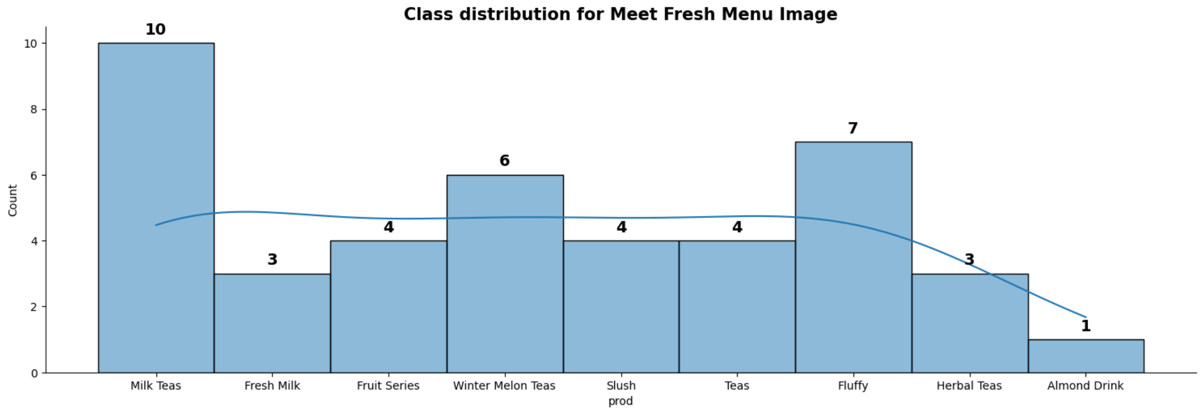

Fig 5: class distribution of MeetFresh Menu Image

- <u>What algorithm to use? And how the algorithm leverage the data.</u>
  For the recommendation engine, the entire workflow of how we develop the engine is shown below:

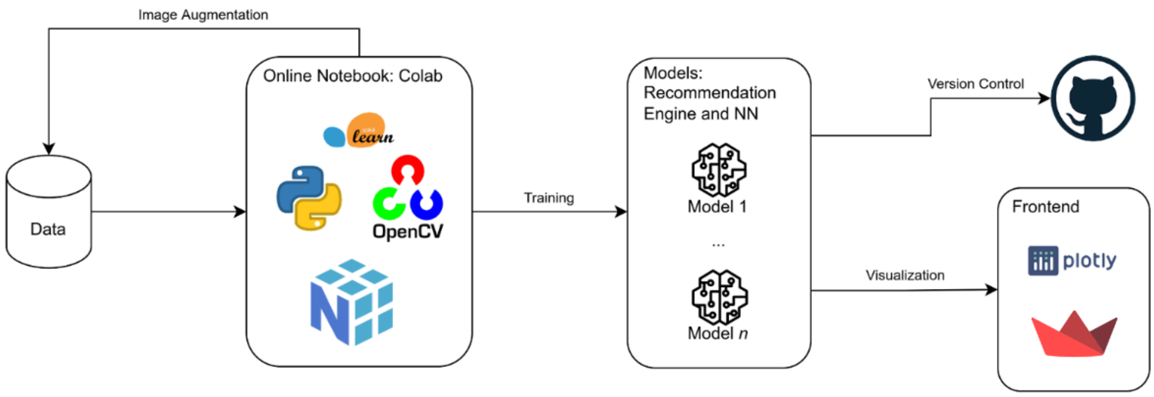

Fig 6: workflow of the entire system development process

Firstly, we utilized cosine similarity score to compute a new user's feature matrix to an existing user's feature matrix based on similarity. The cosine similarity allows us to maximize the use of data by mapping user to user.

For the computer vision part, we utilized Nearest Neighbor (NN) to perform image segmentation. This feature is utilized in the case when the clients have items on their cart, based on the cart items, and recommend other products that are similar to them. In our data preprocessing step, we performed image augmentation by edge detection and cropping since some of the images have a background on them. Performing image augmentation enables us to have subject independent (since we do not count the background as the features for our model) and enrich our training data. Also, over-sampling is performed to balance the image data since the class distribution of the image is not balanced. Moreover, for the feature engineering step, we extract various handcraft features for our NN model, such as Histogram of Oriented Gradients (HoG), Local Binary Pattern (LBP), and Scale-invariant feature transform (SIFT). After training the model with those handcraft features, we return to the participants who completed our survey and let them test the NN model. The result demonstrated the participants enjoyed the NN model that has SIFT features. As a result, we select the SIFT features as our final choice. During the feature engineering, we also tried the feature selection method, Principal Component Analysis (PCA). However, due to the lack of interpretability of the component, we deprecated it during the feature engineering process.

Solution development and results

Based on our data models, we have the distance between each two products from meetfresh. We recommend the product that has the closest distance with the product our customer chooses. For 50% of products, different models recommend the same products. We choose the will For now, we do not have conditions to conduct a big range real world A/B testing, but we do ask 35 people to choose the favorite recommendation model. Scale-invariant feature transform (SIFT) has the best performance compared to other model.But the difference is not significant compared to other models.In order to get more precise results and more proper evaluation, we need more samples and well designed testing.

Summary

Our solution is applied as what we designed, a recommendation engine, and an enhanced recommendation engine, but the detail can be different. We considered using nlp to get reviews for each product but it can be hard to do it while we develop tech solutions. There is too much noise in each review and there are few reviews for new or not popular product.Instead of using NLP, we use different parameters like sensitivity to price, allergen and so on. Based on these parameters and the data we collect for people's preference, we develop a recommendation engine which works well.

For the enhanced recommendation engine, instead of recommending similar products, we decide to recommend the product that has similar looks because how a product looks can be really important for a desert based on our interview with customers and professions.

We also have our design for future development.Once these two engines actually applied to meet fresh stores. We are going to have more samples and evaluation methods.We designed three steps to further increase the accuracy of our model.

The first step is to use A/B testing to see which recommendation model has the best performance. We can randomly assign these models to different customers and count the conversion rate. We can see if the recommendation product was chosen by a customer which helps us to decide which model we are going to use in future.

The Second step is to develop a real time data visualization dashboard to help us to track the performance of the model. We can add KOC rate, false positive rate and other statistical indicators to help us evaluate the performance of our system.

The third step is using the new data we get from customers to train our models repeatedly. The accuracy and performance of the model will become better and better as the data quantity increases. We can also add new products into the system during updating models.