

Yelp Restaurant Attributes Prediction

Yu Wen, Huijing Zhang, Qian Shen, Shiyu Zhang

1. Problem and Goal

Most people rely on Yelp to locate great restaurants, write reviews, and upload restaurant photos to Yelp via their mobile devices. Yelp is now hosting tens of millions of photos shared by Yelpers all over the world. Yelp aims to add attribute labels for each restaurant photo shared by Yelpers, classifying restaurants into different categories. By labeling restaurants with distinct business attributes, Yelp is able to translate uploaded photos into more explicit category information for each restaurant. There are 9 different business attributes listed, 1) good for lunch, 2) good for dinner, 3) takes reservations, 4) outdoor seating, 5) restaurant is expensive, 6) has alcohol, 7) ambience is classy, 8) has table service, 9) good for kids. Currently, these restaurant labels can only be manually selected by Yelp users while submitting a review. Since this selection is optional, many uploaded restaurant photos are not or partially classified. To Yelp users, when looking for certain category restaurants, these restaurant labels will help them quickly find out the ones that satisfy their requirement. For example, Sam would like to treat his parents a great dinner. He will be interested in restaurants whose attribute label is good for dinner. With the help with restaurant category label, he is able to quickly find out desirable restaurants.

There is no doubt that, classifying restaurants into different categories can better serve users' request in a more efficient way. Uploaded restaurant photos themselves contain great information which could be provided to users, helping them make a better decision of whether choosing a restaurant or not. Relying on only Yelp users to label restaurants is undesirable, since some users may forget to select the attribute or are willing to select at all. According to Yelp researchers, there are only a small number of users who would like to category restaurant photos uploaded. However, the accuracy of manually selected features is not reliable. It is entirely possible that, users select a feature at will without careful consideration because of limited time or other issues. However, appropriate classification of these restaurant photos plays a significant role from the website's side perspective. Lacking thorough analysis of attributes for each restaurant photos will lead to poor website performance. In a worse case, users may abandon Yelp if all the restaurants are listed without providing detailed and accurate category information, or their categories are presented in a mass. Thus, the more accurate classification information we can mine from these photos the better.

In this project, we would like to get rid of manually labeling from Yelp users. Instead, develop a data mining technique to build a restaurant photo classifier. This classifier is responsible for taking a restaurant photo as input, and automatically attaching some of the above-mentioned 9 business attributes to it based on the algorithm developed. Our expectation is that, after building the classifier using training data, the classifier is able to predict business attribute for each restaurant photo in an

accurate way. With the help of this classifier, Yelp can gain a more detailed and precise analysis over these uploaded restaurant photos and enhance the quality of service.

2. Data Plan

In this project we will use the image data and the preprocessed data transferred into CSV format. Details are as followings,

- 1) train_photo: an folder contains photos of the training set
contains 234,843 photos of restaurant, each photo has a photo_id
- 2) test_photo: an folder contains photos of the test set
contains 7GB photos of restaurant, each photo has a photo_id
- 3) train_photo_to_biz_ids.csv: maps the train_photo id to business id, for example

| photo_id | business_id |
|----------|-------------|
| 204149 | 3034 |
| 52779 | 2805 |
| 278973 | 485 |

- 4) test_photo_to_biz_ids.csv: maps the test_photo id to business id, similar to train_photo_to_biz_ids.csv
- 5) train: the main training dataset which includes the business id's and their corresponding business labels, which is also the filed we are going to predict, for example

| business_id | labels |
|-------------|---------------|
| 1000 | 1 2 3 4 5 6 7 |
| 1001 | 0 1 6 8 |
| 100 | 1 2 4 5 6 7 |
| 1006 | 1 2 4 5 6 |

Since this project comes from a kaggle competition, above dataset can be download from the website directly.(<https://www.kaggle.com/c/yelp-restaurant-photo-classification/data>)

3. Solution Plan

In our project, we will firstly train the model. In this step, we need the training photos and `train_photo_to_biz_ids.csv`, which maps the photo id to business id, as well as `train.csv` which gives the true labels of the training set. Since the test photos are used for competition submission which didn't contain true labels, so, we will train and validate our model by using cross validation method. Then we will submit our prediction result by using the test photos and `train_photo_to_biz_ids.csv`, in this step we can get the accuracy of our model given by Kaggle.

We may use a feed forward multilayer neural network to build our model and use backpropagation algorithm to optimize it. We may also use the restricted boltzman machine to preprocess our photo, since this algorithm can cluster photos by extract some features automatically. For example, we may want to preprocess our photos by given it some tag such as "outdoor", "indoor" and "alcohol".

4. Evaluation Plan

Since we got data from Kaggle website, they offer a test data set for us. So we will train our model on the training set and predict labels on the test data set, and the performance can be given by submitting the prediction result to kaggle, where we can get our mean F1-Score to exam the accuracy of our model. Meanwhile, we will also use cross-validation method to see our average mean F1-Score on the training

Here is an explanation of our evaluation method. Mean F1-Score, also known as example-based F-measure in the multi-label learning literature. It considers both the precision p and the recall r of the test to compute the score: p is the number of correct positive results divided by the number of all positive results, and r is the number of correct positive results divided by the number of positive results that should have been returned. The F_1 score can be interpreted as a weighted average of the precision and recall, where an F_1 score reaches its best value at 1 and worst at 0.

The traditional F-measure or balanced F-score (**F_1 score**) is the harmonic mean of precision and recall:

$$F_1 = 2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}}$$

The F_1 metric weights recall and precision equally, and a good retrieval algorithm will maximize both precision and recall simultaneously. Thus, moderately good performance on both will be favored over extremely good performance on one and poor performance on the other.

Also, when we need to do classify, we think we can also draw an ROC curve during we are choosing the threshold of our algorithm. ROC analysis provides tools to select possibly optimal models and to discard suboptimal ones independently from (and prior to specifying) the cost context or the class distribution.

5. Schedule

| Date | Milestones |
|------------|--|
| 2015-03-05 | Implementation of tag recognition |
| 2015-03-12 | Acquire an accuracy above the baseline |
| 2015-03-22 | Midterm Report due |
| 2015-04-12 | Submission to Kaggle |
| 2015-04-27 | Final Report & Code due |

Reference

1. <https://www.kaggle.com/c/yelp-restaurant-photo-classification>