Yelp is founded to help people find out great local business, like hair stylists, mechanics, dentists and especially restaurants. 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 one of the above-mentioned 9 business attributes to it based on the algorithm developed. We will use training dataset provided by Yelp to train our restaurant attribute classifier. After the construction of the classifier, we will test it over the test dataset, see what level of the accuracy the classifier can achieve. 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.