**Recommending locations and categories relative to business requirements.**

**Team Members**:

1. Suyog Swami
2. Vineet Nair

**Objective:**

* To recommend business categories for a given location requirement and suggest best Working Hours.
* To recommend locations for a given category input and suggest best area for the business in the suggested location.

**Overview:**

Today setting up a new business or expanding the existing business to new locations is a challenging task. The success of the business depends on many factors like locations proximity from happening places like colleges, offices social gatherings etc. Also popularity (reviews) of the business matters extensively in its success. We try to build a user interactive site where the users can find out best suitable places to their desired businesses. Also we help them find out category of businesses in the specified location which are more likely to succeed.

**Data Mining Tasks:**

We will partition the yelp business dataset into clusters having similars business category or location. We will apply centroid based clustering to identify the target cluster based on users specification. Next task will be grouping the records in the cluster depending on their location or business category. The records in the cluster will be further grouped by their working hours. This will help us to determine the best working hours based on count of records in the group.

**End Goals:**

The end product will be a user interactive website where user will specify the requirements and we will recommend appropriate output (category/location). The output might be in form of visual graphs or a list of outputs in descending order of different criteria.

**Challenges:**

The main challenge will be converting the users input location into its respective longitude and latitude and matching it with the yelp datasets (business) longitude latitude to find closest possible businesses. The second challenge would be to accurately recommend locations for a given business category which are nearby the users specified location. E.g. If user specifies Arlington as the location the resulting categories of businesses should be from nearby cities like Irving, Plano, Dallas etc. The third challenge will be estimating the relation of user input query with the target attributes. Eg. If user enters “schezwan” as business category then we should be able to map “schezwan” with “Chinese Restuarant”.

**Overcoming Challenges:**

For addressing the first challenge, we will use Google’s geocoding API for calculating the longitudes and latitudes. The second challenge can be tackled by applying appropriate clustering techniques to get the locations closest to the input location. For third challenge we will use the attributes in the business dataset which mentions the alternative descriptions for the record.

**Evaluating the Efficacy:**

Using the longitude and latitude of the input location we will find the closest businesses to the location. Then we will use the cluster to find the appropriate businesses in the city depending upon the count of records for particular business category.

Similar execution will be applied to the category wise recommendation to find appropriate location. The closeness of the result with the input provided will help us evaluate the efficacy of the website and its results.

**Task Partitioning:**

Since we have two similar feature implementations in this course project, each team member will implement the respective feature. We will finally work together towards the integration purpose of the project.