CS5560: Knowledge Discovery and Management

First project report

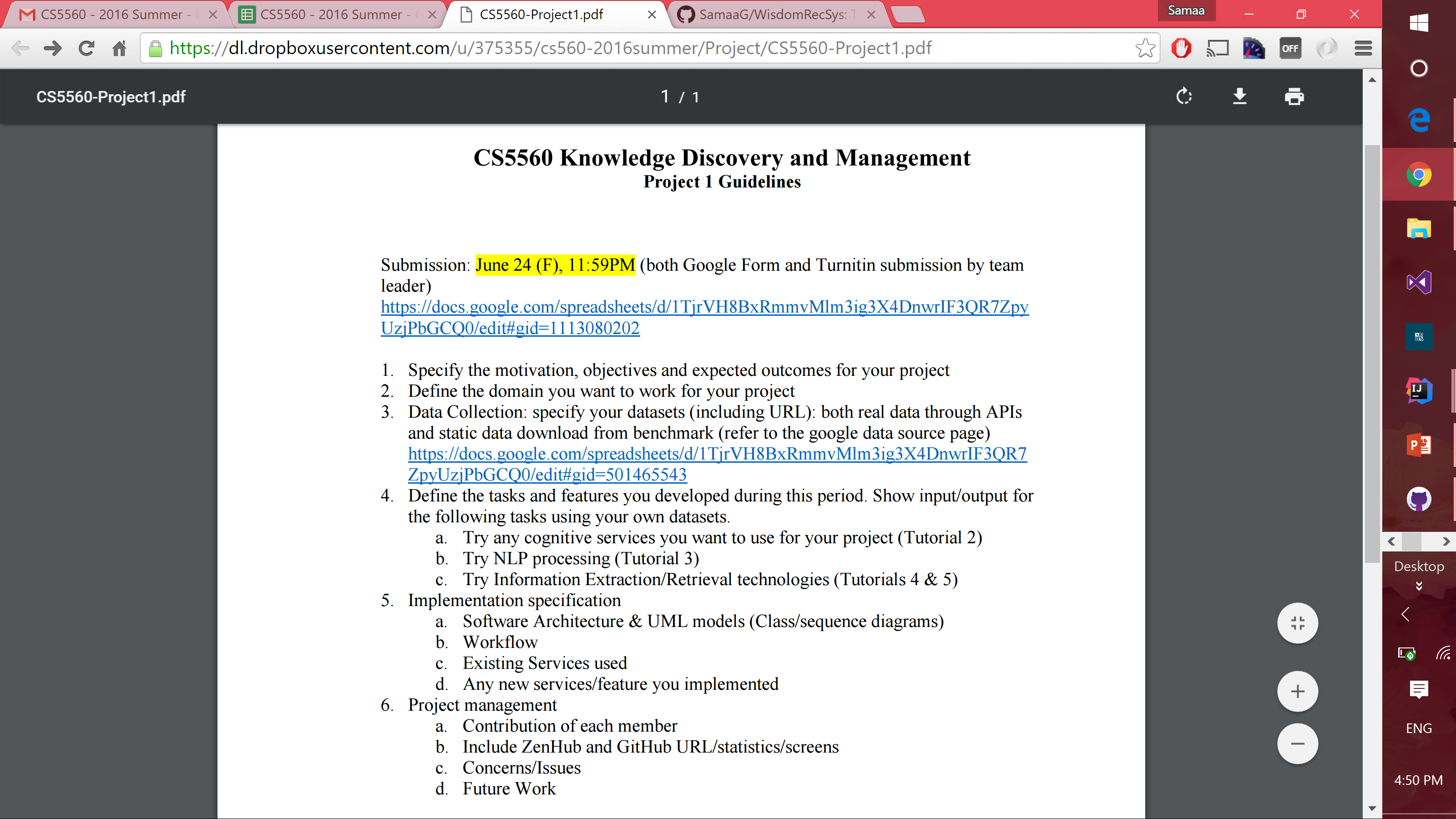
Restaurant recommendation system

Team#5: Wisdom

**Team members**

Samaa Gazzaz (9), Pooja Shekhar (38), Chen Wang (44), Dayu Wang (45)

**This page is just for reference and will be deleted before submission.**



# Motivation

Finding a good restaurant to try new cuisines or even for a fun night out can be a really hard decision to make; especially with all the new restaurants competing to acquire new customers. Using restaurant recommendation apps or systems is very useful but some important features are still missing. When using most of the existing restaurant recommendation systems, recommended restaurant almost never change when searching from the same location-usually your home. This can be a problem as the same users would want to find new restaurants when searching for a restaurant two consecutive days. In addition, users usually have to read a lot of reviews in order to decide whether they found a restaurant that meets their needs. Usually, just looking at the five-star rating is not enough.

# Objective

There are multiple necessary features that are still very much needed in currently existing restaurant recommendation systems. The objective of creating our own restaurant recommendation system is to try and fill the gap by providing those missing features. This system should be able to recommend new restaurants to users that match their needs even if users ask for recommendation from the same location and preferences as a previous search; recommendations should be different. In addition, this system will categories and rank different features of restaurant using sentiment analysis of user reviews. This way, new users don’t have to read all the reviews. Instead, they can just check a visual summery of the features and their ranking for each restaurant.

# Expected Outcome

When we finish developing this system, it should be able to provide restaurant recommendation to users. Users should be able to specify different kind of preferences when asking for a recommendation such as location, type, closing time…etc. Recommended systems should be updated if the user asks for recommendation two consecutive days; not showing the same recommendations. In addition, ranking of different ranking of each recommended restaurant should be provided depending on analyzing the customer reviews. For example, in addition to five-star ratings, a restaurant should also have a ranking of cleanliness, noise level, friendliness and other features users might have mentioned in the reviews.

# Project Domain

This project is going to be a recommendation system. Namely, it will be specialized in recommending restaurants to users. In addition, the specific domain is involved in restaurants and food chains. Not only will it provide the recommendation, our system will make it easy for users to decide where to eat by summarizing existing reviews by categorizing and raking different features mentioned in the reviews.

# Data Collection

## Dataset Used:

For this project, we will be using Yelp restaurant information. Yelp is a hybrid app that provides users with the ability to search for restaurant by specifying their preferences. Then, Yelp provides a list of restaurants which could be reorganized and sorted by rating, nearest location or other preferences.

## Collection Process:

Yelp provides developers with valuable resources and accessibility to static and dynamic datasets. For the static dataset, we will be using the challenge dataset provided by Yelp for developers that want to use their information in research and join a challenge at the same time. That data set is 2.2GB in size and includes 2.2M reviews and 591K tips by 552K users for 77K businesses. It can be accessed and downloaded through the following URL: <https://www.yelp.com/dataset_challenge>

Moreover, Yelp also provides a well-documented API for developers to access the real-time data featured on Yelp. This API provides access to search over 50 million local businesses from 32 countries. The API URL is: <https://www.yelp.com/developers/>

# Tasks and Features

# Input/output

## Cognitive Services:

## NLP Processing:

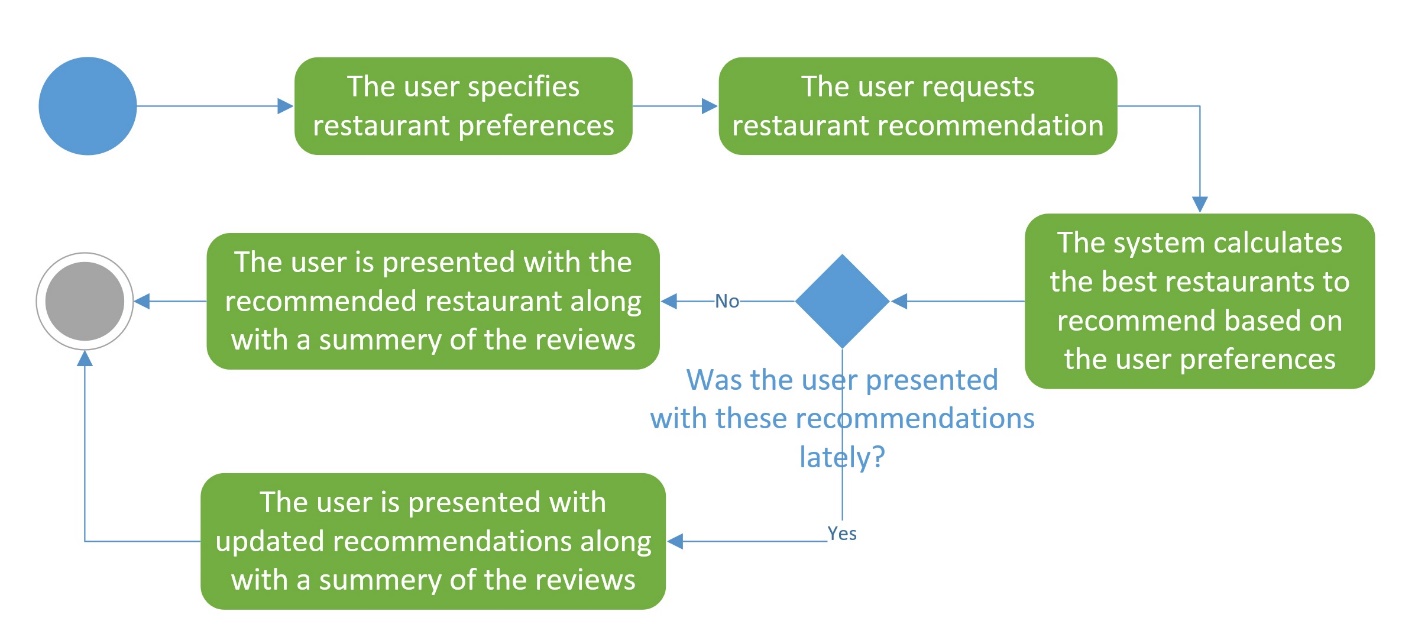
## Information Extraction/Retrieval Technology:

# Implementation Specification

## Software Architecture:

## UML Models:

## Workflow:



## Existing Services:

## New Services/Features Implemented:

# Project Management

## Contribution of Each Member:

|  |  |  |
| --- | --- | --- |
| Member | Contribution | |
| Samaa Gazzaz | Documentation:   * Motivation * Objective * Expected Outcome * Project Domain * Project Management * Workflow Diagram\* | Dataset collection:   * Collecting Static Datasets   Input/Output:   * Information Extraction/Retrieval Technology\*   + Code is included in the repository |
| Pooja Shekhar |  | |
| Chen Wang |  | |
| Dayu Wang |  | |

## Version Control/Screenshots:

For this project, we used GitHub as the main version control tool. The whole project, in addition to documentation, is up on: <https://github.com/SamaaG/WisdomRecSys>

## Concerns/Issues:

Although there are plenty of currently available recommendation systems that could serve as reference for this project, the time restrain imposes a huge concern on whether our system could be implemented on time. Since this is the first exposure for all our team members with recommendation systems, the time needed to get up to speed is going to take away from actual implementation time.

## Future Work:

For the next iteration, we plan to implement at least one feature for the system. In addition, preparing the dataset should be done by the next report.