**CSE 4334/5334 – Data Mining Spring 2015 – Course Project**

**Project Title**

*Application to Determine Trending Restaurants*

**Team participants**

**Rajesh Nayak UTA id: 1001126237**

**Vidyadhar Angadiyavar UTA id: 1001103525**

**Objective and overview of the project**

Determine the trending restaurants by rating them on different categories, considering different attributes such as user reviews, ratings and their respective locations.

We plan to list a number of trending restaurants based on the number of reviews in the past 24 hours and the keywords in the user reviews. Also, we attempt to list the trending restaurants based on ambience, quality of food, cuisine and other categories so that we get a list of different list of trending restaurants for every category based on how a user provides his/her ratings.

We will also find out whether a particular restaurant has moved up or has gone down the trending list from the previous trending data.

**Data Mining Tasks:**

We are planning to use sentiment analysis on the reviews and rate the restaurants on factors such as ambiance, food quality, service, social hangout etc. We also need to determine whether a review is a positive or negative because some of the reviews would sound positive but are sarcastically made.

Also an analysis is made on the reviewer to check whether he is liberal or strict by performing pattern analysis on the total reviews made by him.

**Challenges**

To extract a rating for ambience, quality of food and other factors based on textual reviews of a user will be difficult. Some reviews might be sarcastic one which will be also a challenge to identify them.

Also, a user might be liberal or strict with his ratings. We can analyze the rating history of the user to understand and calculate the weight of his rating when compared to others.

**Solution**

We will applymachine learning methods such as latent semantic analysis, ‘bag of words’ etc. on the reviews to determine whether the review is positive or a negative one, which will also determine some whether a review is a sarcastic one or not.

To determine how the user reviews, patterns are identified and classify the users among the strict, liberal users. A knowledge discovery database is maintained to classify the users.

**Evaluate Efficacy**

From this application we expect to get the list of restaurants that are trending throughout the week.

**Project Roles**

Rajesh will implement the sentiment analysis on the reviews and part of web application. Vidyadhar will implement the pattern analysis on the user determine the type of the reviewer he/her is and remaining part of the web application.