

## Course Project Proposal

Team: Mango

**Project Coordinator:** Saket Vissapragada

**Team Members:**

Ameya Gharpure ([ameyapg2@illinois.edu](mailto:ameyapg2@illinois.edu))

Sruthi Kilari ([skilari2@illinois.edu](mailto:skilari2@illinois.edu))

Saket Vissapragada ([saketv2@illinois.edu](mailto:saketv2@illinois.edu)),

Shreya Sharma ([ssharm90@illinois.edu](mailto:ssharm90@illinois.edu))

Amrith Balachander ([amrithb2@illinois.edu](mailto:amrithb2@illinois.edu))

**Chosen Topic:** Intelligent Browsing - Standardizing Google Restaurant Reviews

**Issue:** A common problem in restaurant reviewing is that there is no standardization behind the stars that are given in a restaurant review. We want to alleviate this issue by using sentiment analysis to give an unbiased and standardized ranking to all the text-based reviews.

**Programming language:** Python/Javascript

**Theme and course relation :** Our project relates to the theme of intelligent browsing by integrating a chrome extension that provides added functionality to Google review. This project relates to the course by using web scraping by parsing HTML files and creating a model.

**Estimated workload:**

- 20 hours \* 5 members = 100 total hours for this project
- Project Planning and Task Assignment
  - 10 hours total = 2 hours/person \* 5 group members
- Web Scraper for obtaining text from Google reviews
  - 25 hours total = 5 hours/person \* 5 group members
- Apply algorithms covered in class to generate sentiment analysis scores for each review based on commonly used words
  - 30 hours total = 6 hours/person \* 5 group members
- Rank these scores against each other to determine standardized star ratings and compute an average standardized score for restaurants.
  - 25 hours total = 5 hours/person \* 5 group members
- User testing to determine accuracy of standardized scores
  - 10 hours total = 2 hours/person \* 5 group members

**Demonstrating approach accuracy:** We will have volunteers assess what they believe the scores of reviews should be. We will present the reviews without the scores and use them to measure the accuracy of the scores generated from our algorithm.