**Capstone Project Submission**

**Instructions:**

i) Please fill in all the required information.

ii) Avoid grammatical errors.

|  |
| --- |
| **Team Member’s Name, Email and Contribution:** |
| **TEAM MEMBERS -**   1. IQBAL BABWANE 2. SAMEER ANSARI 3. LUKMAN HAIDER   **NAME-** SAMEER ANSARI  **EMAIL –** an.sameer.1998@gmail.com  **CONTRIBUTION -**  Data Analysis, Data Visualization, Feature Engineering, Fitting Models, Model Explainability and Report Writing. |
| **Please paste the GitHub Repo link.** |
| https://github.com/reemas-irasna/Zomato-restaurant-clustering-and-sentiment |
| **Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)** |
| * Natural Language Processing is one part of Artificial Intelligence and Machine Learning to make an understanding of the interactions between computers and human (natural) languages. Sentiment analysis is one part of Natural Language Processing, that often used to analyze words based on the patterns of people in writing to find positive, negative, or neutral sentiments. Sentiment analysis is useful for knowing how users like something or not. Zomato is an application for rating restaurants. The rating has a review of the restaurant which can be used for sentiment analysis. Based on this, we want to discuss the sentiment of the review to be predicted. The method used for preprocessing the review is to make all words lowercase, tokenization, remove numbers and punctuation, stop words, and lemmatization. Then after that, we create word to vector with the term frequency-inverse document frequency (TF-IDF). The data that we process are contains some reviews. After that make positive with reviews that have a rating of 3 and above, negative with reviews that have a rating of 3 and below, and neutral who have a rating of 3. After performing Sentiment Analysis, we got some conclusion as follows- * Almost **79** percent of the observations has **positive** sentiment and **14** and **7** percent of the observations has **Neutral** and **Negative** sentiments respectively. * **Good** is the most common word in the highly positive sentiment. * **Worst** is the most common word in the highly negative sentiment. * **Udipi's Upahar** is Affordable Restaurant with best ratings. * **Club Rogue** is Expensive Restaurant with worst ratings. |
|  |