**Capstone Project Submission**

**Instructions:**

i) Please fill in all the required information.

ii) Avoid grammatical errors.

|  |
| --- |
| **Team Member’s Name, Email and Contribution:** |
| Rinkesh Das  Email Id- [rinkeshdas2001@gmail.com](mailto:rinkeshdas2001@gmail.com)  Contribution- Data Cleaning, EDA, Data preprocessing and Feature engineering. Model Training, Evaluation and Validation |
| **Please paste the GitHub Repo link.** |
| Github Link:- https://github.com/rinkeshdas01/Online-Retail-Customer-Segmentation |
| **Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)**  **Problem Statement-** The task is to identify major customer segments on a transnational dataset which contains transactions occurring between 01/12/2010 and 9/12/2011 for a UK based and registered non-store online retail. The company mainly sells unique all occasion gifts.  **This project contains the following steps:-**   1. Data Cleaning 2. Exploratory Data Analysis (EDA) 3. Data Preprocessing 4. Model Training and Implementation 5. Visualization of the clusters formed in various models 6. Model Evaluation and Validation   **After performing the steps mentioned above the following conclusions were obtained:-**   * From the evaluation metrics, KMeans Clustering has the lowest Davies Bouldin Score and highest Calinski Harabasz Score and Silhouette Score. * It is observed that in KMeans Clustering, Cluster 0 contains 41% of the customers , Cluster 1 contains 29% of the customers and the Cluster 2 contains 30% of the customers. So the distribution is quite acceptable. * So, we choose KMeans Clustering as the optimal model for our problem. |
|  |