**Problem Statement:**

Customer Segmentation Using Machine Learning

**Team Members:**

1. Koppuravuri Sruthi Laya
2. Dodda Prathima
3. Gollapalli Akshitha
4. Gunda Manasa
5. Doolla Hari Priya

**Approach of the Solution:**

1. First we imported all the libraries like numpy, pandas, seaborn, matplotlib…
2. We imported the dataset called ‘data2.csv’.
3. We explored the features and checked for the duplicates.
4. We found the word frequencies in the dataset.
5. Next we used KNN algorithm to segment the customers.
6. We used Principle Component Analysis as it is a technique for analysing large data sets containing a high number of dimensions or features and it is used to reduce the number of features.
7. Then we used KNN algorithm as it stores all the available data and classifies a new data point based on the similarity.
8. Therefore we finally segmented the customers.

**Algorithms:**

* Support Vector Machine Classifier (SVC)
* Logistic regression
* k-Nearest Neighbours
* Decision Tree
* Random Forest
* AdaBoost
* Gradient Boosting Classifier
* Principle Component Analysis

**Final Accuracy:**

75.46%