

Abstract:

The goal of this project was to help the manager to predict the right group of new customers by dividing all customers into four groups in the dataset (A, B, C, D). to conduct segmented outreach and communication for a certain customer segment

Design:

This project originates from An automobile company that has plans to enter new markets with their existing products (P1, P2, P3, P4, and P5)., After intensive market research, The company deduced that the behavior of the new market is similar to their existing market.

Data:

This dataset was acquired from the Analytics Vidhya hackathon, provided by Kaggle with 11 Features and contains 425.37 kB

Algorithms:

Data Model:

- 1- LogisticRegression
- 2- RandomForest
- 3- KNeighbors
- 4- DecisionTree

	LogisticRegression	RandomFores	KNeighbors	DecisionTree
Accuracy	46.80%	45.90%	41.86%	42.54%
Recall	46.80%	42.54%	41.86%	45.90%
Precision	46.80%	42.54%	41.86%	45.90%
F1-Score	46.80%	43%	41.86%	45.90%

as we see the Logistic Regression give the best result in our model

Tools:

NumPy | Pandas | seaborn | Matplotlib

Conclusion:

In conclusion, the classifications were used to pick the right group of new clients with whom to conduct awareness and communication.