***Project Report on Marketing Data Prediction***

The dataset includes CSV file of 2240 observations (customers) with 28 variables related to marketing data. More specifically, the variables provide insights about:

* Customer profiles
* Products purchased
* Campaign success (or failure)
* Channel performance

Aim to achieve:

1. Out of the 27 independent variables, variables that are important in predicting ‘MntWines’ column are identified.
2. On these variables, PCA has been applied to reduce the number of independent variables.
3. ML models with created Principal Components to predict ‘MntWines’: Multiple Linear Regression, SVM, Decision Trees.

***PROJECT INFERENCES:***

1. Out of the 27 independent variables, finding variables that are important in predicting ‘MntWines’ column include

NumStorePurchases

NumCatalogPurchases

Income

MntMeatProducts

NumWebPurchases

Kidhome

AcceptedCmp5

2.PCA has been applied to reduce the no of independent variable from 27 to 15.

3.Model predictions:

1. **Linear Regression Model**

Multiple Linear Regression Score: 0.6980059281574307

MSE: 31577.62461315771

RMSE: 177.7009415089231

R2 Score: 0.6980059281574307

1. **Support Vector Machine Model**

SVM Score: 0.2732350488072841

MSE: 75993.24937321001

RMSE: 275.66873122138827

R2 Score: 0.2732350488072841

1. **Decision Tree Regressor Model**

Decision Tree Score: 0.4506862979030728

MSE: 57438.286036036036

RMSE: 239.6628591084485

R2 Score: 0.4506862979030728

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | SCORE | MSE | RMSE | R2 SCORE |
| MLR | 0.6980059 | 31577.6246 | 177.70094 | 0.6980059 |
| SVM | 0.2732350 | 75993.2493 | 275.66873 | 0.2732350 |
| DT | 0.4506862 | 57438.2860 | 239.66285 | 0.4506862 |