
ASSIGNMENT NO.3

Title: Apply appropriate ML algorithm on a dataset collected in a cosmetics shop showing details of customers to predict customer response for special offers.

Software/Libraries Used:

- Python
- pandas
- scikit-learn

Methodology:

1. Data Preprocessing:

- Load the dataset into a DataFrame.
- Handle missing values: Fill or drop missing values.
- Encode categorical variables: Convert categorical variables into numerical format (e.g., one-hot encoding).
- Split the data into features (X) and target variable (y).

2. Model Training:

- Split the data into training and testing sets.
- Initialize and train a classification model (e.g., Logistic Regression) using the training data.

3. Model Evaluation:

- Evaluate the trained model's performance using metrics such as accuracy, precision, recall, or F1-score on the testing data.

4. Prediction:

- Use the trained model to predict customer responses for new data.

Advantages:

- Logistic Regression is a simple and interpretable algorithm.
- It works well for binary classification problems like this.
- Easy to implement and understand.

Limitations:

- Logistic Regression may not capture complex relationships between features.
- It assumes linear decision boundaries, which may not always be appropriate for the data.

Working/Algorithm::

1. Load the dataset and preprocess it.
2. Split the data into features (X) and target variable (y).
3. Split the data into training and testing sets.
4. Initialize and train a Logistic Regression model using the training data.
5. Evaluate the model's performance using metrics such as accuracy, precision, recall, or F1-score.
6. Use the trained model to predict customer responses for new data.

Conclusion:

By applying machine learning algorithms like Logistic Regression to the dataset collected in the cosmetics shop, we can predict customer responses for special offers. While Logistic Regression offers simplicity and interpretability, it's essential to consider other algorithms for potentially better performance, especially for more complex datasets. Additionally, careful evaluation of the model's performance and tuning of parameters can lead to improved predictions.