** Bansilal Ramnath Agarwal Charitable Trust’s**

**Vishwakarma Institute of Information Technology, Pune-48**

**(An Autonomous Institute affiliated to Savitribai Phule Pune University)**

**Department of Computer Science and Engineering (Artificial Intelligence)**

**LAB SUBMISSION**

**Data Science and Machine Learning**

**CAUA22201**

*Submitted by:*

*Avdhoot Nilesh Pimparkar,*

*PRN: 22210178,*

*Roll Number: 281054*

*Second Year*

*Semester II Academic Year 2023-24*

**Assignment: 3**

Aim: To apply appropriate ML algorithm on a dataset collected in a cosmetics shop showing details of

customers to predict customer response for special offers.

Theory:

To predict customer responses for special offers in a cosmetics shop using machine learning (ML), several steps need to be followed.

1. Data collection from the cosmetics shop is crucial, encompassing various customer details and their responses to special offers. We have taken the dataset from Kaggle. Here is the dataset: “<https://www.kaggle.com/datasets/kingabzpro/cosmetics-datasets>”. This dataset is diverse and includes 1472 rows and 11 columns.

2. After data collection, preprocessing is essential to clean and handle missing values, encode categorical variables, and scale numerical features as necessary. Feature engineering comes next, where new features are created or existing ones transformed to enhance predictive power. For instance, deriving features such as ‘Price’ or ‘Rank’ can provide valuable insights.

3. Subsequently, the dataset is split into training and testing sets for model evaluation. Then, an appropriate ML algorithm, such as logistic regression, decision trees, random forests, or gradient boosting, is chosen based on the nature of the problem (classification) and dataset characteristics. The selected algorithm is trained on the training data and evaluated using the testing data. Evaluation metrics like accuracy, precision, recall, F1-score, and ROC-AUC are commonly used to assess model performance. Once satisfied with the model's performance, it can be deployed in a production environment to predict customer responses in real-time. Continuous monitoring and updating of the model are crucial to ensure its effectiveness in predicting customer responses accurately over time.

Here, I have used Random Forest Classifier as algorithm to build model.

Random Forest Classifier:

**Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset.** Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output. **The greater number of trees in the forest leads to higher accuracy and prevents the problem of overfitting.**



Fig. Random Forest Classifier

**Advantages:**

1. Random Forest is capable of performing both Classification and Regression tasks.
2. It is capable of handling large datasets with high dimensionality.
3. It enhances the accuracy of the model and prevents the overfitting issue.

Limitations:

Although random forest can be used for both classification and regression tasks, it is not more suitable for Regression tasks.

Applications:

There are mainly four sectors where Random Forest Algorithm is mostly used:

1. **Banking:** Banking sector mostly uses this algorithm for the identification of loan risk.
2. **Medicine:** With the help of this algorithm, disease trends and risks of the disease can be identified.
3. **Land Use:** We can identify the areas of similar land use by this algorithm.
4. **Marketing:** Marketing trends can be identified using this algorithm.

Results:

These are the results that are obtained:

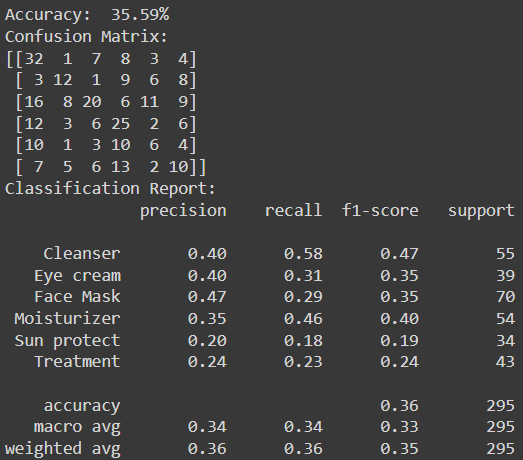
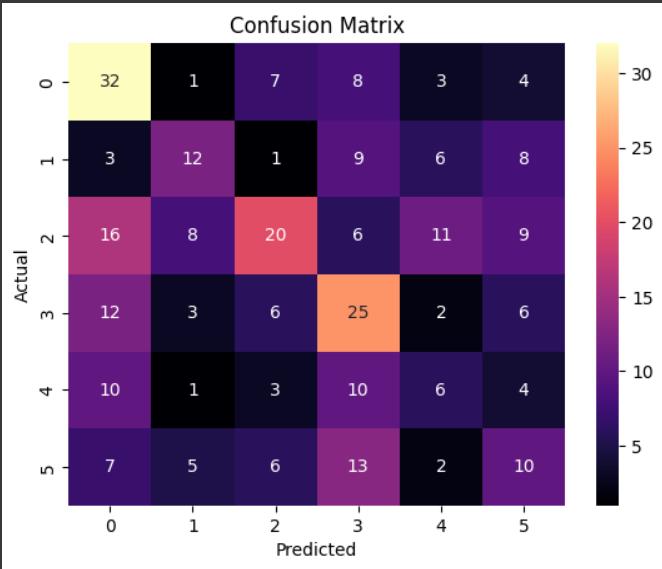
 

Fig. Classification Report Fig. Confusion Matrix

Conclusion:

In this assignment, we were able to implement Random Forest Classifier algorithm on a dataset collected in a cosmetics shop showing details of customers to predict customer response for special offers.