****

**Assessment Report**

on

**“CUSTOMER BEHAVIOR”**

submitted as partial fulfillment for the award of

**BACHELOR OF TECHNOLOGY**

**DEGREE**

SESSION 2024-25

in

**CSE(AI)**

By

Name : Aditya Singh

Roll Number : 202401100300021

Section: A

**Under the supervision of**

“Mr. BIKKI KUMAR”

**KIET Group of Institutions, Ghaziabad**

**May, 2025**

1. **Introduction**

**Understanding customer behavior is key to personalized marketing and improved user experience in e-commerce and retail platforms. This project focuses on classifying customers into two categories: 'bargain hunters' and 'premium buyers', based on their purchase history. This classification allows businesses to target their strategies effectively.**

1. **Problem Statement**

**To classify customers as either 'bargain hunters' or 'premium buyers' using their purchase history features such as total spent, average purchase value, and visits per month. The goal is to identify customer segments for tailored marketing and service strategies.**

1. **Objectives**

* **Preprocess the dataset for training a classification model.**
* **Train a Random Forest classifier to predict customer types.**
* **Evaluate the model using accuracy, precision, recall, and classification report.**
* **Visualize the confusion matrix using a heatmap for interpretability.**

1. **Methodology**

**Data Collection: The user uploads a CSV file containing customer behavior data.**

**Data Preprocessing:**

* **Label encoding of the target variable ('buyer\_type').**
* **Feature selection: total\_spent, avg\_purchase\_value, visits\_per\_month.**
* **Train-test split with 80-20 ratio.**

**Model Building:**

* **Random Forest Classifier is used for its robustness and ensemble nature.**

**Model Evaluation:**

* **Calculate accuracy, precision, and recall.**
* **Generate and visualize the confusion matrix with Seaborn heatmap.**

1. **Data Preprocessing**

* **The 'buyer\_type' is encoded as 0 for 'bargain\_hunter' and 1 for 'premium\_buyer'.**
* **Features used are: 'total\_spent', 'avg\_purchase\_value', and 'visits\_per\_month'.**
* **The dataset is split into 80% training and 20% testing.**

1. **Model Implementation**

**A Random Forest classifier with 100 estimators and fixed random state is trained on the preprocessed dataset. It is then used to predict the buyer type on the test set.**

1. **Evaluation Metrics**

* **Accuracy: Measures the percentage of correctly classified customers.**
* **Precision: Proportion of predicted premium buyers that are actually premium buyers.**
* **Recall: Proportion of actual premium buyers that were correctly identified.**
* **Confusion Matrix: Visualized using a Seaborn heatmap.**

1. **Results and Analysis**

* **The classifier showed good performance in categorizing customers.**
* **Confusion matrix heatmap provides insight into true positives, false positives, true negatives, and false negatives.**
* **Precision and recall indicate the model's effectiveness in distinguishing between the two customer types.**

1. **Conclusion**

**The Random Forest model effectively classifies customers into 'bargain hunters' and 'premium buyers' using basic purchasing behavior data. This classification supports better marketing strategies and user engagement. Future improvements may include more behavioral features and advanced models.**

1. **References**

* **scikit-learn documentation**
* **pandas documentation**
* **Seaborn visualization library**
* **Research on customer segmentation and behavior prediction**





