**Summary Report**

**On**

**Customer Churn Prediction Analysis**



Submitted By:

**Sumit M. Raut**

**Introduction**

**Objective**

The primary goal of this analysis was to predict customer churn and derive actionable insights to improve customer retention and increase sales. The dataset contained customer information, purchasing behaviour, and churn indicators, enabling exploratory data analysis (EDA), predictive modeling, and the identification of key trends and patterns.

**Dataset Overview**

The dataset contains the following fields:

* Customer ID: Unique identifier for each customer.
* Age: Age of the customer.
* Gender: Gender of the customer.
* Annual Income: Customer's yearly income.
* Spending Score: A metric representing spending habits.
* Purchase History: Aggregated count or value of purchases.
* Membership Duration: Duration of the customer's membership in years.
* Feedback Score: Customer satisfaction rating.
* Churn Indicator: Target variable (0 = No, 1 = Yes).

**Approach**

**Data Preparation**

1. Exploratory Data Analysis (EDA):
   * Analyzed data distributions, trends, and anomalies.
   * Identified patterns like correlations between features and churn.
2. Data Cleaning:
   * Handled missing values using median imputation.
   * Removed duplicates and normalized numeric fields.
3. Feature Engineering:
   * Created new features such as **Income-to-Spending Ratio** and **Normalized Purchase History**.
   * Categorized Age into age groups for better segmentation.

**Findings**

**Key Findings from Exploratory Data Analysis:**

**Demographic Patterns**:

* Younger customers (18-25) have a higher churn rate compared to older age groups.
* Older customers exhibit higher loyalty and lower churn risk.

**Behavioural Trends**:

* Shorter membership durations are strongly associated with higher churn rates.
* Low feedback scores correlate significantly with churn, highlighting dissatisfaction as a major driver.

**Spending Patterns**:

* Customers with a high income-to-spending ratio tend to churn more, suggesting less engagement with products or services.
* Spending intensity (spending relative to membership duration) is lower among churned customers.

**Predictive Modeling**

**Models Used:**

1. **Logistic Regression**
2. **Random Forest Classifier**
3. **XGBoost Classifier** (optional model for further exploration)

**Evaluation Metrics:**

* Accuracy
* Precision
* Recall
* F1-score
* AUC-ROC

**Results:**

* **Random Forest** emerged as the most robust model, offering high precision and recall, along with strong feature interpretability.
* **Feedback Score**, **Membership Duration**, and **Spending Intensity** were identified as top predictors of churn.

**Visual Insights**

1. **Churn Rate by Age Group**:

Younger customers exhibit significantly higher churn rates, requiring targeted engagement strategies.

1. **Top Predictors of Churn** (Random Forest Feature Importance):

Membership Duration, Feedback Score, and Spending Intensity are the most critical features influencing churn predictions.

1. **Feedback Score Distribution**:

Churned customers have consistently lower feedback scores, underscoring the importance of addressing dissatisfaction.

**Recommendations**

**Retention Strategies:**

1. **Enhance Onboarding for New Customers**:

Introduce personalized onboarding programs and immediate incentives for new customers to engage them early.

1. **Address Feedback Concerns**:

Develop a proactive feedback resolution team to handle dissatisfaction promptly and boost feedback scores.

1. **Increase Engagement Among Younger Customers**:

Implement gamified loyalty programs, exclusive rewards, and discounts to retain this high-churn demographic.

1. **Reward Loyalty**:

Offer tiered rewards or long-term incentives to customers with extended memberships to reinforce loyalty.

1. **Improve Value Perception for Price-Sensitive Customers**:

Introduce tiered pricing or budget-friendly options to attract and retain low-income customers.

**Revenue Enhancement:**

1. **Personalized Promotions**:

Use insights from the churn model to design tailored marketing campaigns targeting at-risk customers.

1. **Cross-Selling and Upselling**:

Recommend complementary products or premium services based on purchase history and spending behaviour.

1. **Incentivize Feedback**:

Encourage customers to provide positive feedback by offering rewards or recognition, improving overall satisfaction.

**Next Steps**

**Deploy Predictive Model**:

Use the Random Forest model to identify high-risk customers and trigger retention interventions.

**Monitor and Iterate**:

Continuously evaluate the model’s performance and update it with new data to ensure accuracy.

**Action Plan Implementation**:

Roll out the recommended strategies, monitor customer responses, and adjust actions based on outcomes.

**Conclusion**

This analysis highlights the importance of customer satisfaction, tailored engagement strategies, and demographic-specific retention plans. By leveraging the model’s predictions, businesses can proactively reduce churn and drive growth.