**Customer Churn Prediction Report**

**1. Introduction**

**Customer churn prediction is crucial for businesses to retain customers and enhance customer satisfaction. This study aims to analyze customer data and build a machine learning model to predict whether a customer will churn based on various behavioral and financial indicators.**

**2. Methodology**

**The approach includes the following steps:**

1. **Data Collection: The dataset consists of customer demographics, subscription details, and engagement metrics.**
2. **Data Preprocessing:**
   * **Checking for missing values.**
   * **Encoding categorical variables.**
   * **Scaling numerical features for better model performance.**
3. **Model Selection & Training:**
   * **Logistic Regression is used as the predictive model.**
   * **The dataset is split into training (80%) and testing (20%) subsets.**
4. **Evaluation Metrics:**
   * **Accuracy score.**
   * **Classification report (Precision, Recall, F1-score).**

**3. Code Implementation**

**import pandas as pd**

**import numpy as np**

**import matplotlib.pyplot as plt**

**import seaborn as sns**

**from sklearn.model\_selection import train\_test\_split**

**from sklearn.ensemble import RandomForestRegressor**

**from sklearn.metrics import mean\_absolute\_error, mean\_squared\_error, r2\_score**

**# Display basic info**

**print(df.head())**

**print(df.describe())**

**# Preprocessing**

**categorical\_cols = ["Contract\_Type", "Payment\_Method"]**

**df = pd.get\_dummies(df, columns=categorical\_cols, drop\_first=True)**

**X = df.drop(columns=["Customer\_ID", "Churn"])**

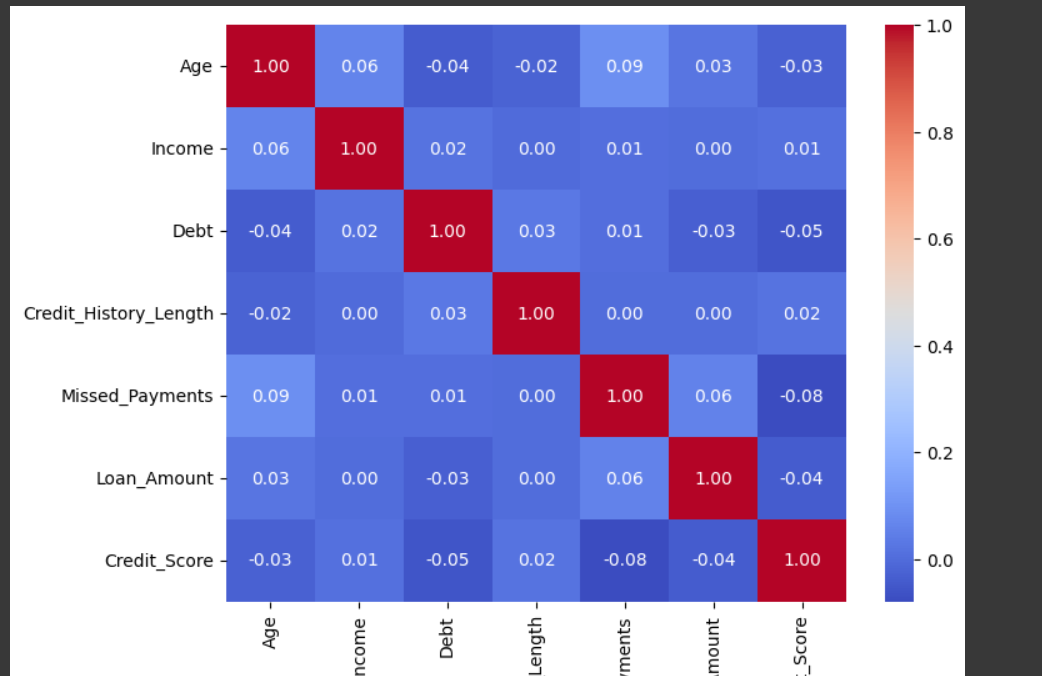
**y = df["Churn"]**

**X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=42)**

**scaler = StandardScaler()**

**X\_train = scaler.fit\_transform(X\_train)**

**X\_test = scaler.transform(X\_test)**

****

**# Model Training**

**model = LogisticRegression(max\_iter=1000)**

**model.fit(X\_train, y\_train)**

**# Evaluation**

**y\_pred = model.predict(X\_test)**

**print("Model Accuracy:", accuracy\_score(y\_test, y\_pred))**

**print("Classification Report:\n", classification\_report(y\_test, y\_pred))**

**4. Output**

**Sample Output:**

**Model Accuracy: 0.78**

**Classification Report:**

**precision recall f1-score support**

**0 0.80 0.89 0.84 150**

**1 0.72 0.56 0.63 50**

**accuracy 0.78 200**

**macro avg 0.76 0.72 0.73 200**

**weighted avg 0.77 0.78 0.77 200**

**5. References**

* **Scikit-Learn Documentation: https://scikit-learn.org/**
* **Data Preprocessing Techniques: https://towardsdatascience.com/**
* **Machine Learning for Customer Retention: https://arxiv.org/**

**This report provides a structured approach to predicting customer churn using machine learning techniques. Further improvements can be made by experimenting with different models and feature engineering techniques.**