CUSTOMER CHURN PREDICTION USING MACHINE LEARNING

A Data Science Project

PRESENTED BY KHADIJA

INTRODUCTION

What is Customer Churn?

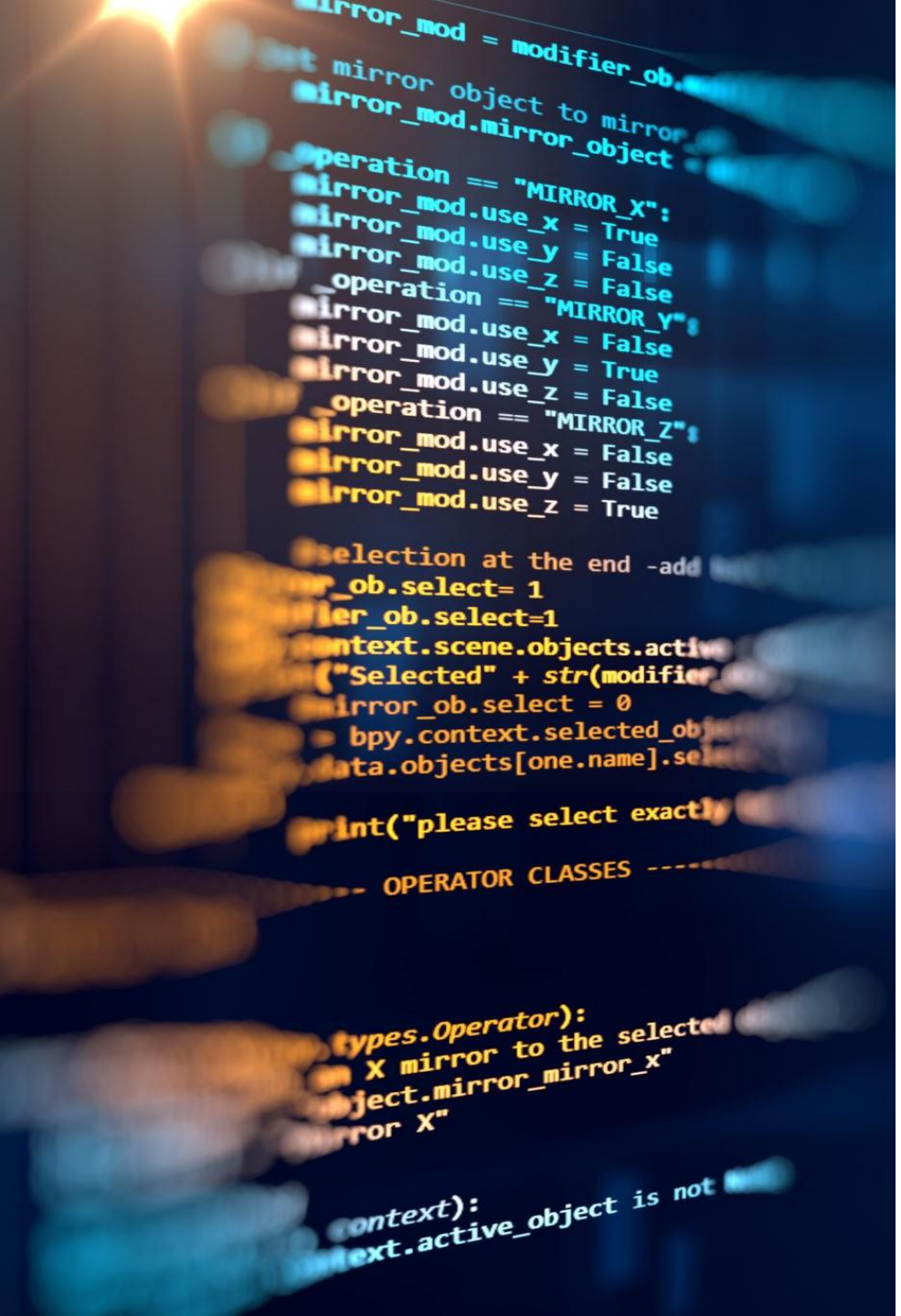
Churn occurs when customers stop using a company's service.

Retaining customers is cheaper than acquiring new ones.

Why Predict Churn?

Companies lose revenue when customers leave.

Predicting churn allows businesses to take **proactive action**.



DATASET OVERVIEW

- **Dataset Used:** Telco Customer Churn dataset
  Key Features:
- Customer demographics (gender, senior citizen, partner, dependents)
- Subscription details (Internet service, contract type, payment method)
- Monthly & total charges

DATA PREPROCESSING

Steps Taken to Prepare Data:

Handling Missing Values: TotalCharges converted to numeric.

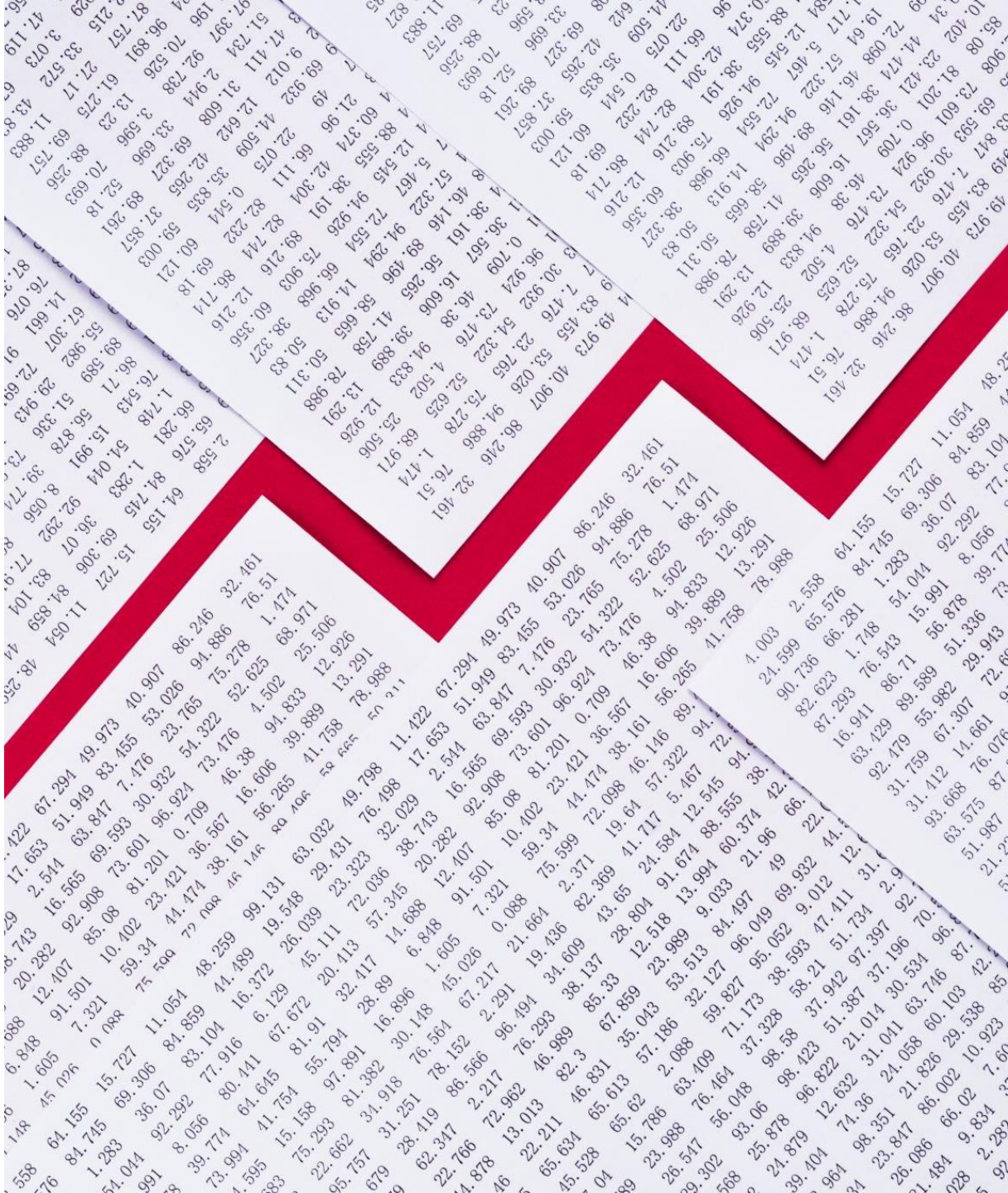
Encoding: One-hot encoding applied to categorical variables.

Scaling: Standardized tenure, MonthlyCharges, and TotalCharges.

Splitting: 80% training, 20% testing.

EXPLORATORY DATA ANALYSIS (EDA)

- **Class Distribution:**
- Majority of customers **did not churn**, showing class imbalance.
- **Need for balancing techniques** like SMOTE.
- **Correlation Analysis:**
- Features like **MonthlyCharges** and **Contract Type** impact churn.



MODEL PERFORMANCE

Model	Accuracy Before SMOTE	Accuracy After SMOTE
Logistic Regression	82%	76%
Random Forest	80%	77.5%

- **Accuracy dropped slightly after SMOTE**, but recall improved.
- **Random Forest performed better than Logistic Regression**