

## PROJECT REPORT

### Customer Churn Prediction Using Logistic Regression

---

#### 1. Introduction

Customer churn refers to the situation where customers stop using a company's services. In the telecom industry, customer churn is a major challenge because acquiring new customers is more expensive than retaining existing ones.

Machine Learning techniques can help organizations predict customer churn in advance and take preventive actions.

This project uses **Logistic Regression**, a supervised machine learning algorithm, to predict whether a telecom customer is likely to churn or not.

---

#### 2. Problem Statement

The objective of this project is to build a machine learning model that can predict whether a customer will leave the telecom service based on customer demographics, services used, and billing information.

##### Target Variable:

- Churn → Yes / No
- 

#### 3. Dataset Description

- **Dataset Name:** Telco Customer Churn Dataset
- **Source:** Kaggle
- **Total Records:** 7043
- **Number of Features:** 20
- **Target Column:** Churn

##### Key Features:

- Customer demographics: gender, SeniorCitizen, Partner, Dependents
- Services: InternetService, OnlineSecurity, TechSupport, StreamingTV
- Account information: tenure, Contract, PaymentMethod
- Charges: MonthlyCharges, TotalCharges

---

## 4. Methodology

The project was implemented using the following steps:

1. Loading the dataset
2. Data cleaning and handling missing values
3. Encoding categorical variables into numerical format
4. Feature scaling using StandardScaler
5. Splitting the dataset into training and testing sets
6. Training the Logistic Regression model
7. Evaluating the model using performance metrics

---

## 5. Model Used

**Logistic Regression** was chosen because:

- The target variable is binary (Churn: Yes/No)
- It is simple, efficient, and interpretable
- It performs well for classification problems

---

## 6. Model Evaluation and Results

### Accuracy

The model achieved an **accuracy of approximately 80%**, indicating that it correctly predicts churn status for most customers.

### Confusion Matrix

[[921 114]

[169 205]]