Customer Churn Analysis – Final Report

1. Introduction

Customer churn is a major challenge in the telecom industry, directly affecting revenue and growth. Retaining customers is more cost-effective than acquiring new ones. This project aims to predict customer churn and provide actionable strategies for retention, using machine learning and data analytics.

2. Abstract

The project involved analyzing telecom customer data to identify patterns leading to churn. A Random Forest Classifier was built to predict churn probability. Model explainability tools (SHAP and ELI5) were used to understand the influence of key features. Customers were segmented into Loyal, At Risk, and Dormant, enabling targeted retention strategies. Key findings highlight that monthly charges, tenure, contract type, and service usage are the main drivers of churn.

3. Tools Used

- Python Libraries: Pandas, NumPy, Matplotlib, Seaborn
- Machine Learning: Scikit-learn (Random Forest Classifier)
- Explainability: SHAP, ELI5
- Environment: Google Colab / Jupyter Notebook

4. Steps Involved in Building the Project

1. **Data Loading & Preprocessing:**
    - Loaded customer dataset
    - Converted necessary columns to numeric types
    - Handled missing values and checked data quality

2. **Exploratory Data Analysis (EDA):**
    - Visualized churn distribution
    - Analyzed key features like Monthly Charges, Tenure, Contract type
    - Observed patterns leading to higher churn
3. **Feature Encoding & Model Training:**
    - Encoded categorical features
    - Split data into training and testing sets
    - Built Random Forest Classifier and evaluated model performance
4. **Model Evaluation:**
    - Assessed accuracy and confusion matrix
    - Identified top features influencing churn using SHAP/ELI5
5. **Customer Segmentation:**
    - Categorized customers into Loyal, At Risk, and Dormant
    - Provided segment-specific insights for retention strategies
6. **Recommendations:**
    - At Risk: Personalized campaigns, discounts, proactive outreach
    - Dormant: Engagement programs, promotional offers
    - Loyal: Reward programs, upsell opportunities, maintain satisfaction

5. Conclusion

The project successfully predicted telecom customer churn and identified key factors contributing to churn. Customer segmentation enables targeted retention strategies, improving customer satisfaction and reducing revenue loss. Implementing these recommendations can help telecom companies retain high-value customers and strengthen their competitive position.