# Phase3\_project

# SyriaTel Customer Churn Prediction

This project aims to analyze customer churn behavior for SyriaTel, a telecommunications company, using machine learning techniques. By understanding the factors contributing to customer churn, we can provide actionable insights to reduce customer attrition and improve customer retention strategies.

## Table of Contents

1. [Project Overview](#project-overview)

2. [Dataset Description](#dataset-description)

3. [Methodology](#methodology)

4. [Technologies Used](#technologies-used)

5. [Results and Insights](#results-and-insights)

6. [How to Use](#how-to-use)

7. [Contributing](#contributing)

---

## Project Overview

Customer churn refers to the loss of customers over a given period. For telecommunications companies like SyriaTel, minimizing churn is critical for maintaining profitability. This project:

- Uses machine learning models to predict churn probability for customers.

- Identifies key factors influencing customer churn.

- Provides recommendations to reduce churn.

## Dataset Description

The dataset used in this project contains information about SyriaTel's customers, including:

- Demographics: Age, gender, etc.

- Service usage: Internet service, streaming services, etc.

- Churn label: Indicator of whether the customer has churned (binary).

### Key Features:

- `customerID`: Unique identifier for each customer.

- `MonthlyCharges`: The amount charged to the customer monthly.

- `Churn`: Target variable indicating whether the customer churned.

## Methodology

1. \*\*Data Cleaning and Preprocessing:\*\*

- Handle missing values.

- Convert categorical variables into numerical representations.

- Standardize numerical features.

2. \*\*Exploratory Data Analysis (EDA):\*\*

- Visualize distributions and relationships between features.

- Identify patterns and trends influencing churn.

3. \*\*Model Training:\*\*

- Comparison of multiple models including Decision Trees, Random Forests, and Logistic Regression.

- Use hyperparameter tuning (e.g., Grid Search) to optimize models.

4. \*\*Evaluation Metrics:\*\*

- Accuracy

- Precision

- Recall

- F1-score

- ROC Curve and AUC

5. \*\*Insights and Recommendations:\*\*

- Analyze feature importance to understand churn drivers.

- Propose strategies for customer retention.

## Technologies Used

- \*\*Programming Language:\*\* Python

- \*\*Libraries:\*\*

- Pandas, NumPy: Data manipulation

- Matplotlib, Seaborn: Visualization

- Scikit-learn: Machine learning

## Results and Insights

- \*\*Best Model:\*\* Decision Tree with an accuracy of 92%.

- \*\*Key Factors Influencing Churn:\*\*

- Contract type

- Monthly charges

- Tenure

### Recommendations:

- Offer incentives for long-term contracts.

- Provide competitive pricing for high-risk customers.

- Focus retention efforts on customers in their first year of tenure.

## How to Use

1. Clone this repository:

```bash

git clone https://github.com/your-repo/syriatel-churn-prediction.git

```

2. Install required libraries:

```bash

pip install -r requirements.txt

```

3. Run the Jupyter Notebook for analysis and model training:

```bash

jupyter notebook SyriaTel\_Churn\_Analysis.ipynb

```

## Contributing

Contributions are welcome! If you want to improve this project:

1. Fork the repository.

2. Create a new branch.

3. Submit a pull request.

## License

This project is licensed under the [MIT License](LICENSE).