1. Project Title

Customer Churn Prediction for a Telecom Company

---

2. Business Problem

Customer churn is a major challenge for telecom companies. Losing existing customers directly impacts revenue.
The goal of this project is to **predict whether a customer will churn** so that the business can take preventive actions.

Business Requirement

- The model must be **explainable**
- Predictions should be understandable by non-technical stakeholders

---

3. Dataset Description

The dataset represents customer information from a telecom company.

Features

Feature Name	Description
tenure	Number of months customer stayed
monthly_charges	Monthly bill amount
total_charges	Total charges paid
contract_type	Contract duration
internet_service	Type of internet service
payment_method	Mode of payment
tech_support	Whether tech support is available
churn (Target)	1 = Churn, 0 = Not Churn

Target Variable

- **churn** (Binary Classification)

---

4. Data Preprocessing

- Categorical variables encoded using **Label Encoding**
- Data split into:
    - 75% Training

    - 25% Testing
- No missing values in dataset

---

5. Models Implemented

As required, models from different families were implemented:

5.1 Base & Linear Model – Logistic Regression

- Highly explainable
- Coefficients indicate feature importance
- Used as baseline model

5.2 Ensemble Model – Random Forest

- Combines multiple decision trees
- Improves accuracy
- Provides feature importance

5.3 Boosting Model – XGBoost

- Sequential learning
- Best performance
- Used ROC-AUC for evaluation

---

6. Model Evaluation Metrics

- Accuracy
- Precision
- Recall
- F1-score
- **ROC-AUC (Primary Metric)**

---

7. Model Comparison

Model	**ROC-AUC**
Logistic Regression	~0.78
Random Forest	~0.85
XGBoost	~0.89

---

8. Best Model Selection

XGBoost performed best based on ROC-AUC.
However, **Logistic Regression** is preferred when explainability is critical.

Final Recommendation

- Use **Logistic Regression** for business explanation
- Use **XGBoost** for operational prediction

---

9. Business Insights

- Customers with **month-to-month contracts**
- High **monthly charges**
- No **tech support**

are more likely to churn.

---

10. Conclusion

The project successfully implemented models from multiple ML families.
The solution balances **performance and interpretability**, meeting business requirements.