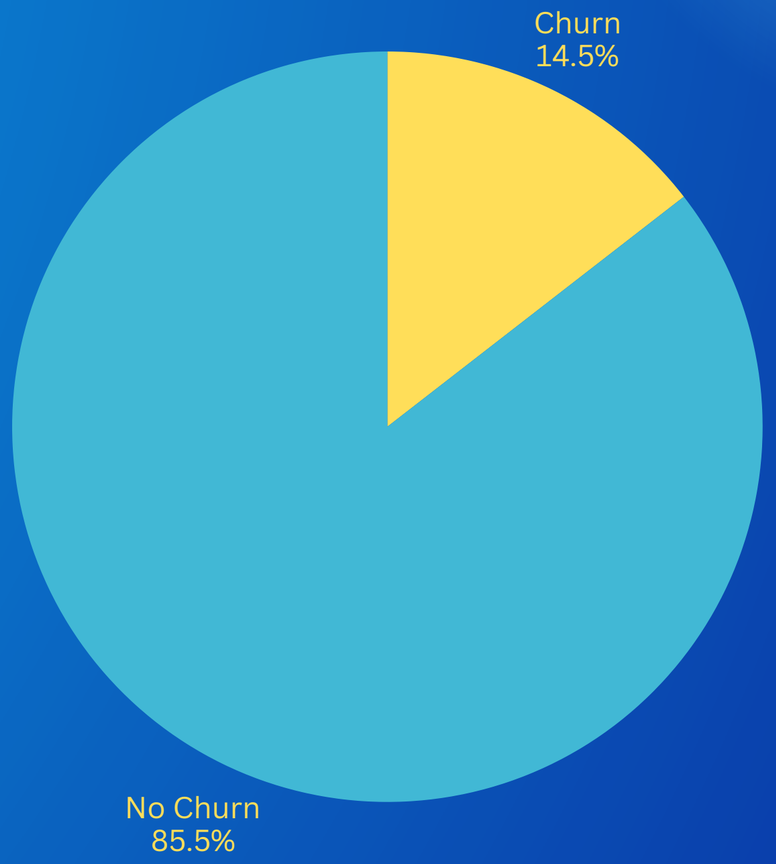
SyriaTel Limited

Problem Statement

- Customer churn is a major challenge in telecommunication companies.
- Losing customers means **lost revenue** and **high replacement costs** to the company.
- **Goal**: Predict which customers are likely to churn and act early.

Dataset Overview

- Data from telecom customers.
- **Includes**: usage patterns, contract details, billing information and demographics.
- ~3,300 records analyzed.
- **Target**: Churn (**Yes** / **No**).

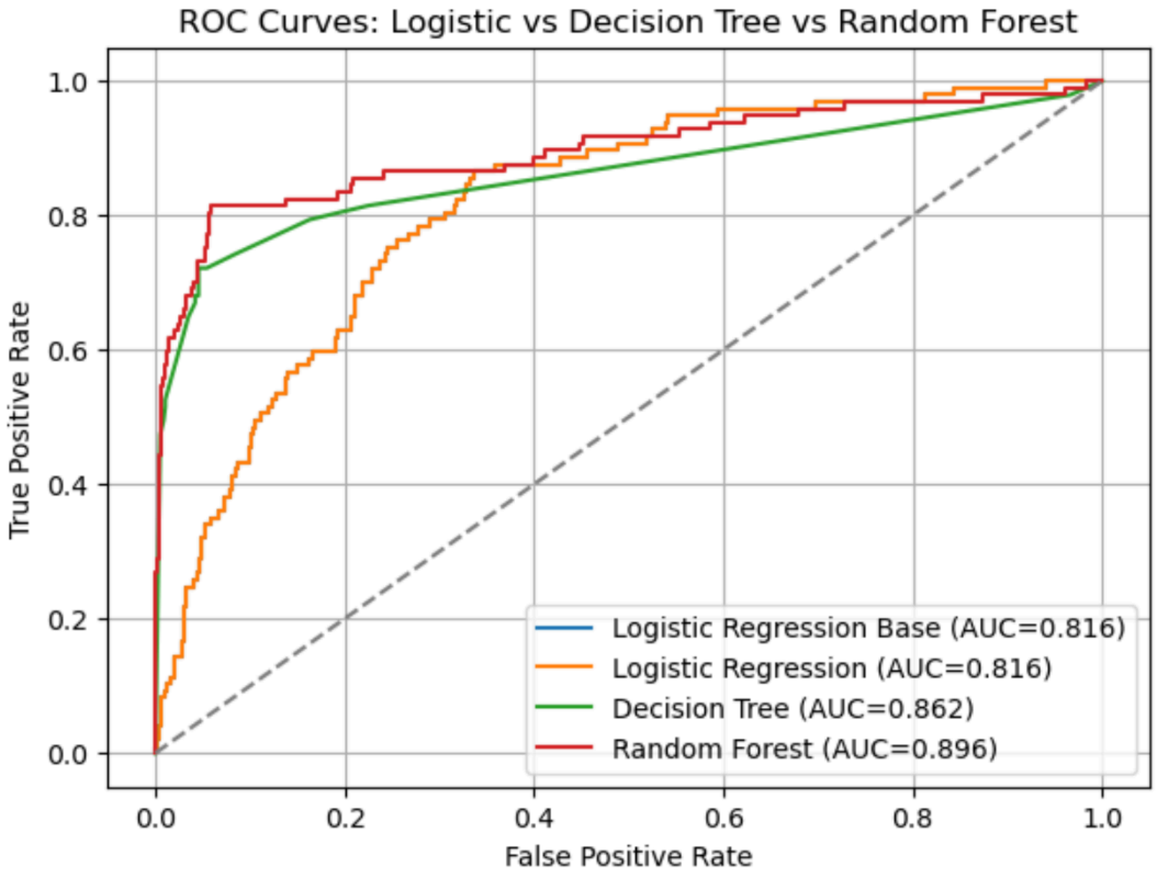


Approach (CRISP-DM)

- **Business Understanding** → Why churn matters.
- **Data Understanding & Preparation** → Cleaning and the exploring the dataset.
- **Modeling** → Used: Logistic Regression, Decision Tree and Random Forest.
- **Evaluation** → ROC AUC comparison.
- **Recommendations** → Use insights for customer retention.

Model Results

- Logistic Regression (Base): 0.815
- Logistic Regression (Tuned): 0.815
- Decision Tree: 0.889
- **Random Forest**: 0.921 **(Best model)**



“MACHINE LEARNING TURNS DATA INTO DECISIONS AND DECISIONS INTO INSIGHT.”

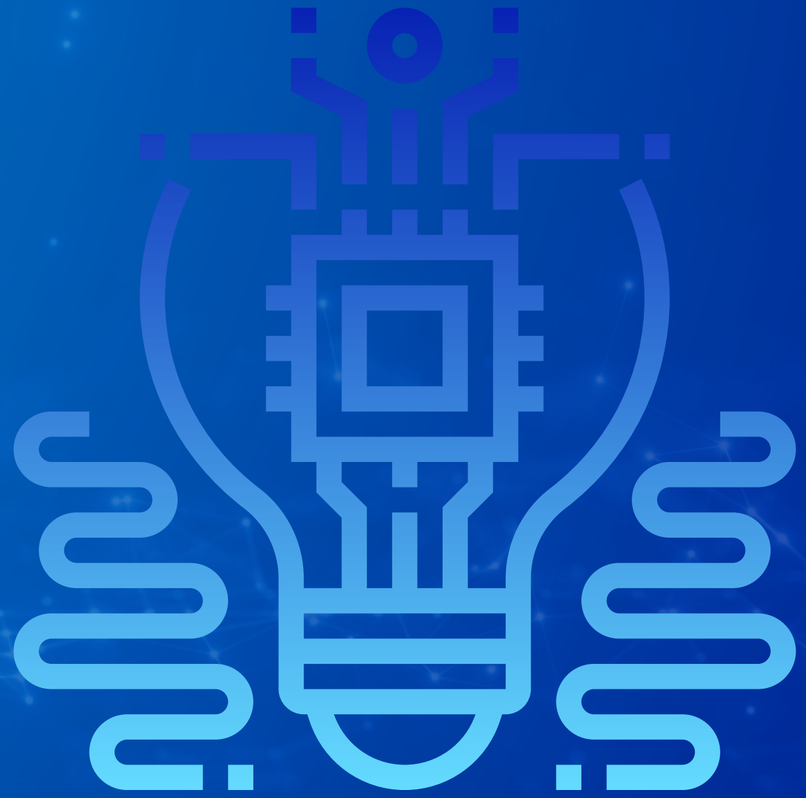
Logistic Regresion(Base) | Logistic Regression | Decision Tree | Random Forest

Insights

- Random Forest is best predicts churn.
- Important churn factors:
  - Contract type (month-to-month vs. long-term).
  - Monthly charges.
  - Customer tenure (length of service).

Recommendations

- Deploy Random Forest for churn prediction.
- Focus on retention efforts on high-risk customers.
- Offer targeted loyalty programs, bundles or discounts.
- Continuously update the model with new customer data.



Business Impact

- **Lower churn** means higher customer retention.
- It saves money on acquiring new customers.
- It improves long-term customer loyalty and revenue stability.

Conclusion

- Data science provides actionable insights for telecom.
- **Random Forest** outperformed other models.

- **Next step**: Deploy the model and integrate with customer support systems.