# SyriaTel Customer churn Analysis

# Outline

- Business Problem
- Data used for Analysis
- Methods for Classification
- Evaluation
- Conclusion
- Recommendation

# Business Problem

SyriaTel telecommunication company faces a significant challenge in managing customer churn, which affects revenue stability and market competitiveness. Despite various retention strategies, the churn rate remains a potential risk, indicating a need for a more data-driven approach to understand and mitigate churn.

# Data Understanding

To gain a comprehensive understanding of the data obtained from SyriaTel, facilitating the identification of key patterns and insights that will inform the development of an effective customer churn predictive model.

The dataset provided by SyriaTel contains detailed information on customer demographics, service usage, billing, and customer service interactions. This understanding phase aims to thoroughly explore and comprehend the structure, quality, and nuances of the data to ensure accurate analysis and modeling.

# Data Preprocessing

Steps to ensure the data  is ready for Customer Churn analysis the following steps were followed;

- **Data Cleaning**: Address any identified data quality issues through cleaning and preprocessing steps, it includes the following;

    Check for null values

    Check Duplicates

- **Model Preparation**: Prepare the data for modeling by splitting it into training and test sets, normalizing numerical attributes, and encoding categorical variables.

# Models Modelling

## Baseline Logistic Regression Model Performance:

Training Accuracy: 89.4%

Test Accuracy:86.0%

**After Hyperparameter Tuning:**

Training Accuracy:90.4%

Test Accuracy:88.8%

## Conclusion:

Hyperparameter tuning improved the training accuracy by 1% and the test accuracy by 2.8%, demonstrating an overall enhancement in the model's performance and generalizability.

# Decision Tree Classifier

**Decision Tree Classifier Model Performance**:

Training Accuracy:90.4%

Test Accuracy:92.5%

**After Hyperparameter Tuning:**

Training Accuracy:90.4%

Test Accuracy:93.6%

**Conclusion:**

Hyperparameter tuning maintained the training accuracy at 90.4% and improved the test accuracy by 1.1%, indicating a significant boost in the model's performance on unseen data.

# Random Forest

**Random Forest Model Performance:**

Training Accuracy**:** 87.1%

Test Accuracy**:** 87.2%

**After Hyperparameter Tuning:**

Training Accuracy**:** 93.6%

Test Accuracy**:** 92.7%

**Conclusion:**

Hyperparameter tuning significantly improved the training accuracy by 6.5% and the test accuracy by 5.5%, indicating a substantial enhancement in the model's overall performance and generalization ability.

# Conclusion

- All three models showed improvements in performance after hyperparameter tuning.
- The Random Forest model displayed the most substantial improvement, making it a strong candidate for predicting customer churn in the SyriaTel dataset.
- Consistent improvements in test accuracy across models suggest that hyperparameter tuning is an effective strategy for enhancing model performance in this analysis.