Name: Rana Emad Tawfik

**Problem:** Understand why customers are leaving and develop predictive models to identify customer churn.

**Goal:** To identify factors affecting customer churn and develop strategies to retain customers back on the next three months.

**Solution:**

**Business Understanding:**

The company’s success will be measured by:

Reduction in churn rate within the next three months.

High accuracy and interpretability of the model to support retention strategies.

What customer factors are most included with churn?

How can we classify customers as high-risk for churn?

What strategies can this company implement for at-risk customers?

**Data Understanding:**

Analyze the dataset containing customer demographics, account details, services used, and churn status Exploration to the following:

Visualization: Use histograms, box plots, and correlation matrices to identify trends and correlations.

Churn Rate: Determine the baseline churn rate to understand how big is the problem.

Descriptive Statistics: Summarize data using mean, median, and standard deviation.

Churn Analysis: Compare features between churned and non-churned customers to identify significant differences.

**Data Preparation:**

Data Cleaning: Address any missing or incorrect values, replace missing values using median or mode, depending on the feature type

Feature Transformation and normalization:

transform categorical variables e.g., contract type into numerical form using one-hot encoding or label encoding. standardize numerical data e.g., monthly charges to ensure features are on a similar scale for modeling.

**Modeling:**

Choose few models suitable for classification problems to predict churn risk maybe by

Regression- Random Forest-Decision tree

Divide the data into training and testing sets to test the model’s predictive capability.

Use Random Search to optimize parameters for each model to improve accuracy.

Apply cross-validation to validate the model and avoid overfitting.

**Evaluation:**Use relevant metrics to assess model performance in identifying churn:

Accuracy – F1 Score – R2 Score – Mean Squared Error

Based on evaluation results, select the model with the highest performance metrics and interpretability for deployment.

**Deployment:**

Integrate the model into company’s customer management system, allowing churn risk assessment.

Implement a system that flags high-risk customers for proactive outreach by customer service teams.

Based on the model’s predictions, develop personalized retention strategies:

consider offering discounts to switch to annual contracts.

Reward long-term customers who may be at risk of churning.

Collect feedback from customers

track model accuracy and retrain with updated data as churn rates may change over time.