**CONNECTTEL CHURN PROJECT**

**Problem Definition**: Predicting Customer Churn and Proactively Winning Them Back.

**Introduction:** Unveiling the Enigma of Customer Churn: Understanding, Anticipating, and Conquering the Attrition Challenge. In the bustling landscape of modern business, where customer loyalty is the cornerstone of success, there exists a formidable adversary: customer churn. The silent yet relentless force that stealthily erodes profitability and undermines growth, customer churn remains an enigma that haunts businesses across industries. It is the unsettling departure of once-loyal patrons, the gradual fading of revenue streams, and the elusive puzzle that demands urgent attention.

In this dynamic and hypercompetitive marketplace, where customer expectations evolve at the speed of innovation, understanding the intricate dynamics of churn is paramount. Every churned customer represents not just a lost revenue opportunity, but a trove of invaluable insights waiting to be unearthed. Behind every departure lies a story—a tale of unmet expectations, unsatisfied needs, or unaddressed grievances. It is within these narratives that the keys to unlocking customer retention reside.

At the intersection of data science, behavioral psychology, and business strategy lies the promise of unraveling the mysteries of customer churn. By leveraging the power of advanced analytics, predictive modeling, and prescriptive insights, businesses can transcend the realm of mere speculation and proactively anticipate churn triggers. Armed with this foresight, they can deploy targeted interventions, personalized incentives, and strategic initiatives to stem the tide of attrition and foster enduring customer relationships.

**Background:**

ConnectTel telecommunication Company is a leading provider of mobile communication and internet service in the Telecommunication sector. Retaining customers is crucial for the company's long-term success and profitability. However, the company has been experiencing a high rate of customer churn, where a significant number of customers discontinue their subscription or stop using the company's services.

**Objective:**

The objective of this project is to develop a predictive model that can identify customers who are at risk of churning. By proactively identifying these customers, the company aims to implement targeted retention strategies to reduce churn rates and improve customer retention. In the context of confusion matrix, any model or group of models that will **maximize the true positive** (Knowing ahead of time more of the customers that likely to churn) and **minimize the false negative** (Having less of people we erroneously assumed to be loyal but are on their way out) is our best choice.

**Data:**

The company has collected historical data on customer interactions, including demographic information, usage patterns, transaction history, customer tenure, type of contract, Monthly charges, customer support interactions, and other relevant features. The dataset includes both churned and active customers, with labels indicating whether each customer churned or not. Find below a detailed data dictionary:

**DATA DICTIONARY**

**1. CustomerID**: A unique identifier assigned to each telecom customer, enabling

tracking and identification of individual customers.

2. **Gender**: The gender of the customer, which can be categorized as male, or

female. This information helps in analyzing gender-based trends in

customer churn.

3. **SeniorCitizen**: A binary indicator that identifies whether the customer is a senior citizen

or not. This attribute helps in understanding if there are any specific

churn patterns among senior customers.

4. **Partner**: Indicates whether the customer has a partner or not. This attribute helps

in evaluating the impact of having a partner on churn behavior.

5. **Dependents**: Indicates whether the customer has dependents or not. This attribute

helps in assessing the influence of having dependents on customer

churn.

6. **Tenure**: The duration for which the customer has been subscribed to the telecom

service. It represents the loyalty or longevity of the customer’s

relationship with the company and is a significant predictor of churn.

7. **PhoneService**: Indicates whether the customer has a phone service or not. This attribute

helps in understanding the impact of phone service on churn.

DATA DICTIONARY

8. **MultipleLines**: Indicates whether the customer has multiple lines or not. This attribute helps in analyzing

the effect of having multiple lines on customer churn.

9. **InternetService**: Indicates the type of internet service subscribed by the customer, such as DSL, fiber optic,

or no internet service. It helps in evaluating the relationship between internet service and

churn.

10. **OnlineSecurity**: Indicates whether the customer has online security services or not. This attribute helps in

analyzing the impact of online security on customer churn.

11. **OnlineBackup**: Indicates whether the customer has online backup services or not. This attribute helps in

evaluating the impact of online backup on churn behavior.

12. **DeviceProtection**: Indicates whether the customer has device protection services or not. This attribute helps

in understanding the influence of device protection on churn.

13. **TechSupport**: Indicates whether the customer has technical support services or not. This attribute helps

in assessing the impact of tech support on churn behavior.

14. **StreamingTV**: Indicates whether the customer has streaming TV services or not. This attribute helps in

evaluating the impact of streaming TV on customer churn.

15. **StreamingMovies**: Indicates whether the customer has streaming movie services or not. This attribute helps in understanding the influence

of streaming movies on churn behavior.

16. **Contract**: Indicates the type of contract the customer has, such as a month-to-month, one-year, or two-year contract. It is a crucial

factor in predicting churn as different contract lengths may have varying impacts on customer loyalty.

17. **PaperlessBilling**: Indicates whether the customer has opted for paperless billing or not. This attribute helps in analyzing the effect of

paperless billing on customer churn.

18. **PaymentMethod**: Indicates the method of payment used by the customer, such as electronic checks, mailed checks, bank transfers, or credit

cards. This attribute helps in evaluating the impact of payment methods on churn.

19. **MonthlyCharges**: The amount charged to the customer on a monthly basis. It helps in understanding the relationship between monthly

charges and churn behavior.

20. **TotalCharges**: The total amount charged to the customer over the entire tenure. It represents the cumulative revenue generated from the

customer and may have an impact on churn.

21. **Churn**: The target variable indicates whether the customer has churned (canceled the service) or not. It is the main variable to

predict in telecom customer churn analysis.

**Detail of Key Tasks performed**:

This is a step-by-step process flow for a machine learning-based churn problem project, from data cleaning to model deployment:

1.Data Collection: Gather historical customer data from various sources, such as databases, CRM systems, or transaction logs.

2. Data Cleaning:

- Handle missing values,

- Handle outliers

- Check for duplicates

- Standardize data formats

3. Feature Engineering

- Create new features

- Encode categorical variables: Convert categorical variables into numerical format using techniques like Minmax encoding

4. Exploratory Data Analysis (EDA)

- Analyze data distribution, trends or patter

- Visualize relationships using various plots (e.g., histograms, scatter plots) to visualize relationships between variables and identify potential correlations with churn.

- Identify key features: Identify features that may have a significant impact on churn based on their distribution and relationship with the target variable.

5. Data Splitting

- Split the dataset into training and test sets: with 80%:20% respectively.

6. Model Training

- Choose appropriate algorithms: Select machine learning algorithms suitable for the churn prediction task, such as logistic regression, decision trees, random forests, or gradient boosting machines.

- Train multiple models: Train several models using different algorithms to compare their performance.

- Evaluate performance: Assess model performance using appropriate evaluation metrics (F1-score, ROC AUC)

7. Hyperparameter Tuning:

- Perform hyperparameter optimization:

- Tune model parameters.

8. Feature Importance Analysis:

- Analyze feature importance: Determine the relative importance of each feature in predicting churn using techniques like permutation importance, SHAP values, or coefficients from linear models.

9. Model Evaluation:

- Evaluate final model performance: Assess the performance of the best-performing model(s) on the test set to ensure generalization to unseen data.

10. Model Deployment:

- Deploy the trained model: Integrate the final model into the production environment, making it available for making real-time predictions.

- Monitor model performance: Implement monitoring systems to track model performance over time and ensure its effectiveness in predicting churn.

- Update model as needed: Periodically retrain the model using new data and update it with fresh insights to maintain its accuracy and relevance.

This step-by-step process flow provides a structured approach to addressing a machine learning-based churn problem, from data cleaning and feature engineering to model deployment and monitoring. Adjustments may be made based on specific project requirements and domain expertise.

**Success Criteria:**

The success of this project will be evaluated based on the following criteria:

- Development of a predictive model with high accuracy in identifying ***customers at risk of churn***.

- Implementation of targeted retention strategies resulting in a reduction in churn rates.

- Continuous monitoring and improvement of the model to adapt to changing customer behavior and market dynamics.

**Stakeholders:**

- Marketing and Sales Teams: Utilize churn predictions to implement targeted marketing campaigns and retention offers.

- Customer Service Teams: Proactively reach out to at-risk customers to address their concerns and improve satisfaction.

- Executive Management: Monitor overall churn metrics and assess the impact of retention strategies on business performance.

**Deliverables:**

- Predictive model for customer churn with documentation on model architecture, training process, and evaluation metrics.

- Insights into key factors driving churn behavior and recommendations for retention strategies.

- Deployment of the model into production for real-time churn prediction and monitoring.