**Project Proposal: Predictive Analysis of Customer Churn in the Telecommunications Industry**

**Executive Summary:** In the highly competitive telecommunications industry, understanding and addressing customer churn is critical for maintaining market position and profitability. This project aims to analyze and predict customer churn, focusing on identifying high-risk customers and understanding churn patterns to implement effective retention strategies.

**Objective:** To build a predictive model that accurately identifies customers at risk of churn, using a comprehensive dataset that includes customer demographics, service subscriptions, and transaction histories. The model aims to assist telecom companies in reducing churn rates by proactively targeting retention efforts.

**Data Overview:** We will use the following key data points from the provided dataset:

* Customer demographics: Gender, Age, Senior Citizen status, Marital status, Dependents.
* Service details: Phone service, Internet service, Security, Backup, Device Protection, Premium Tech Support, Streaming services.
* Contract and billing information: Contract type, Paperless billing, Payment method, Monthly and total charges, Tenure.
* Churn status: Whether the customer left the company in the given quarter.

**Approach:**

1. **Data Exploration and Preprocessing**: Analyze the dataset to understand customer demographics, service preferences, and churn patterns. Clean and preprocess the data to handle missing values and categorize features.
2. **Feature Engineering**: Create new features that may influence churn, such as customer loyalty indicators and service utilization patterns.
3. **Model Selection and Training**: Evaluate and select appropriate machine learning models for a classification problem, considering the imbalanced nature of the target variable (Churn).
4. **Model Evaluation**: Use the Recall metric for model evaluation due to the imbalanced data. Other metrics like precision, F1-score, and ROC-AUC will also be considered for a comprehensive evaluation.
5. **Insights and Recommendations**: Provide insights into customer behavior patterns related to churn and suggest targeted retention strategies.

**Expected Outcomes:**

* A robust predictive model for identifying customers at high risk of churn.
* Detailed insights into factors contributing to customer churn.
* Strategic recommendations for targeted customer retention initiatives.

**Data Resourse:**

1. **CustomerID**: A unique ID that identifies each customer.
2. **Gender**: The customer’s gender: Male, Female
3. **Age**: The customer’s current age, in years, at the time the fiscal quarter ended.
4. **Senior Citizen**: Indicates if the customer is 65 or older: Yes, No
5. **Married (Partner)**: Indicates if the customer is married: Yes, No
6. **Dependents**: Indicates if the customer lives with any dependents: Yes, No. Dependents could be children, parents, grandparents, etc.
7. **Number of Dependents**: Indicates the number of dependents that live with the customer.
8. **Phone Service**: Indicates if the customer subscribes to home phone service with the company: Yes, No
9. **Multiple Lines**: Indicates if the customer subscribes to multiple telephone lines with the company: Yes, No
10. **Internet Service**: Indicates if the customer subscribes to Internet service with the company: No, DSL, Fiber Optic, Cable.
11. **Online Security**: Indicates if the customer subscribes to an additional online security service provided by the company: Yes, No
12. **Online Backup**: Indicates if the customer subscribes to an additional online backup service provided by the company: Yes, No
13. **Device Protection Plan**: Indicates if the customer subscribes to an additional device protection plan for their Internet equipment provided by the company: Yes, No
14. **Premium Tech Support**: Indicates if the customer subscribes to an additional technical support plan from the company with reduced wait times: Yes, No
15. **Streaming TV**: Indicates if the customer uses their Internet service to stream television programing from a third party provider: Yes, No. The company does not charge an additional fee for this service.
16. **Streaming Movies**: Indicates if the customer uses their Internet service to stream movies from a third party provider: Yes, No. The company does not charge an additional fee for this service.
17. **Contract**: Indicates the customer’s current contract type: Month-to-Month, One Year, Two Year.
18. **Paperless Billing**: Indicates if the customer has chosen paperless billing: Yes, No
19. **Payment Method**: Indicates how the customer pays their bill: Bank Withdrawal, Credit Card, Mailed Check
20. **Monthly Charge**: Indicates the customer’s current total monthly charge for all their services from the company.
21. **Total Charges**: Indicates the customer’s total charges, calculated to the end of the quarter specified above.
22. **Tenure**: Indicates the total amount of months that the customer has been with the company.
23. **Churn**: Yes = the customer left the company this quarter. No = the customer remained with the company. Directly related to Churn Value.

Reference: <https://www.kaggle.com/blastchar/telco-customer-churn>

<https://community.ibm.com/community/user/businessanalytics/blogs/steven-macko/2019/07/11/telco-customer-churn-1113>