**ENTRANCE TEST (No.2)**

*Position:* ***Data Analyst Intern***

#### Problem Statement

***Customer churn prediction*** is a critical aspect of business management, particularly for industries like telecommunications, internet service providers, pay TV companies, insurance firms, and alarm monitoring services. It involves understanding and addressing customer attrition, which refers to the loss of clients or customers.

For businesses in these sectors, measuring customer attrition is a vital business metric. This is because retaining an existing customer is significantly more cost-effective than acquiring a new one. As a result, thes companies often have customer service branches dedicated to re-engaging customers who are considering leaving. This is because the long-term value of recovered customers far outweighs that of newly acquired ones.

***To address customer churn, predictive analytics comes into play***, churn prediction models to assess the likelihood of customers leaving. These models prioritize a small list of potential defectors, enabling businesses to concentrate their customer retention efforts on those who are most at risk of churning.

#### Key Questions:

* 1. What is the significance of Churn Rate for stakeholders (Customers, MCI, etc.)? (1pt)

***Churn rate is a critical metric for stakeholders as it indicates the percentage of customers who stop using the company's services over a given period. High churn rates can indicate customer dissatisfaction, poor service quality, or better offers from competitors. For MCI, understanding churn rate helps in identifying areas of improvement and implementing strategies to retain customers.***

* 1. What are the characteristics of each Type of Customer (Churn or Not Churn)? (2.5pt)

***1. Distribution of the target variable (Churn):***

***- The distribution chart of the target variable (Churn) shows that the number of customers who did not churn (Non-Churned) is higher than the number of customers who churned (Churned). This indicates that the churn rate in this dataset is quite low.***

***2. Analysis of categorical features:***

***- State: The distribution chart of the `State` variable by `Churn` shows that the churn rate may vary between states. Some states have higher churn rates compared to others.***

***- International plan: The distribution chart of the `International plan` variable by `Churn` shows that customers with an international plan have a higher churn rate compared to those without an international plan.***

***- Voice mail plan: The distribution chart of the `Voice mail plan` variable by `Churn` shows that customers with a voice mail plan have a lower churn rate compared to those without a voice mail plan.***

***3. Analysis of numerical features:***

***- Total day minutes: Customers with higher total day minutes tend to churn more.***

***- Customer service calls: Customers with a higher number of customer service calls tend to churn more. This may indicate that customers are dissatisfied with the service and therefore tend to leave.***

* 1. Which ML modeling can be implemented and represent model results? including features input and explaining features important. (4.5pt)
  2. What actions regarding qualitative and quantitative analytics could be implemented to enhance retention rate? (2pt)

This analysis aims to leverage comprehensive customer data to improve customer engagement and satisfaction for MCI's offerings.

1. **Requirements:**

* Clear, simple, optimized code (if possible) + code explanation. (OOP is a plus but optional)
* Slide is clear, insightful and delightful (optional)
* English is a must (If slide and code are not presented in English, your result will be eliminated)
* Please show Step by Step training modeling in Appendix

**Keyword:** Binary Classification, Python, EDA, Statistical Significancei

#### Reference:

Priority on clarifying insights and descriptive statistics. You can refer to the way statistics are described in these articles:

* [Home Credit Complete EDA & Feature Importance](https://www.kaggle.com/code/codename007/home-credit-complete-eda-feature-importance)
* [What Determines Price of a Laptop?](https://www.kaggle.com/code/michau96/what-determines-price-of-a-laptop)
* [HR Analytics Prediction - Why do people resign?](https://www.kaggle.com/code/paramarthasengupta/hr-analytics-prediction-why-do-people-resign#Is-income-the-main-factor-towards-employee-attrition?)