**Instructions**

1. Install the necessary packages
2. Import you data and perform basic data exploration phase
   * Display general information about the dataset
   * Create a pandas profiling reports to gain insights into the dataset
   * Handle Missing and corrupted values
   * Remove duplicates, if they exist
   * Handle outliers, if they exist
   * Encode categorical features
3. Based on the previous data exploration train and test a machine learning classifier
4. Create a streamlit application (locally)
5. Add input fields for your features and a validation button at the end of the form
6. Import your ML model into the streamlit application and start making predictions given the provided features values

**Note:**

1. This checkpoint should be done locally, (Google colab notebooks won't work)
2. Make sure to run the app using the "streamlit run" command in your terminal.

**Understanding the problem :**

**Churn** in a customer dataset refers to the phenomenon where customers stop doing business with a company or service. It’s often used in the context of subscription services, telecommunications, banking, and other industries where companies track customer retention and loss.

**Key Aspects of Churn:**

1. **Definition:** Customer churn is typically defined as the percentage of customers who discontinue their relationship with a company during a given time period.
2. **Types of Churn:**
   * **Voluntary Churn:** When customers choose to leave for reasons such as dissatisfaction, better offers elsewhere, or changing needs.
   * **Involuntary Churn:** When customers are lost due to factors beyond their control, such as payment issues or service discontinuation.
3. **Churn Rate:** This is a critical metric for businesses, calculated as:

Churn Rate=Number of Customers LostTotal Number of Customers at the Start of the Period×100\text{Churn Rate} = \frac{\text{Number of Customers Lost}}{\text{Total Number of Customers at the Start of the Period}} \times 100Churn Rate=Total Number of Customers at the Start of the PeriodNumber of Customers Lost​×100

This metric helps businesses understand how well they are retaining customers.

1. **Predicting Churn:** Data analysis techniques, including machine learning, can be used to predict which customers are likely to churn. Common features used in churn analysis include:
   * Customer demographics (age, location)
   * Purchase history
   * Customer service interactions
   * Subscription duration
   * Usage patterns
2. **Importance of Churn Analysis:**
   * Helps identify at-risk customers so that retention strategies can be implemented.
   * Guides improvements in product or service offerings based on customer feedback.
   * Aids in forecasting future revenue by understanding customer lifetime value.