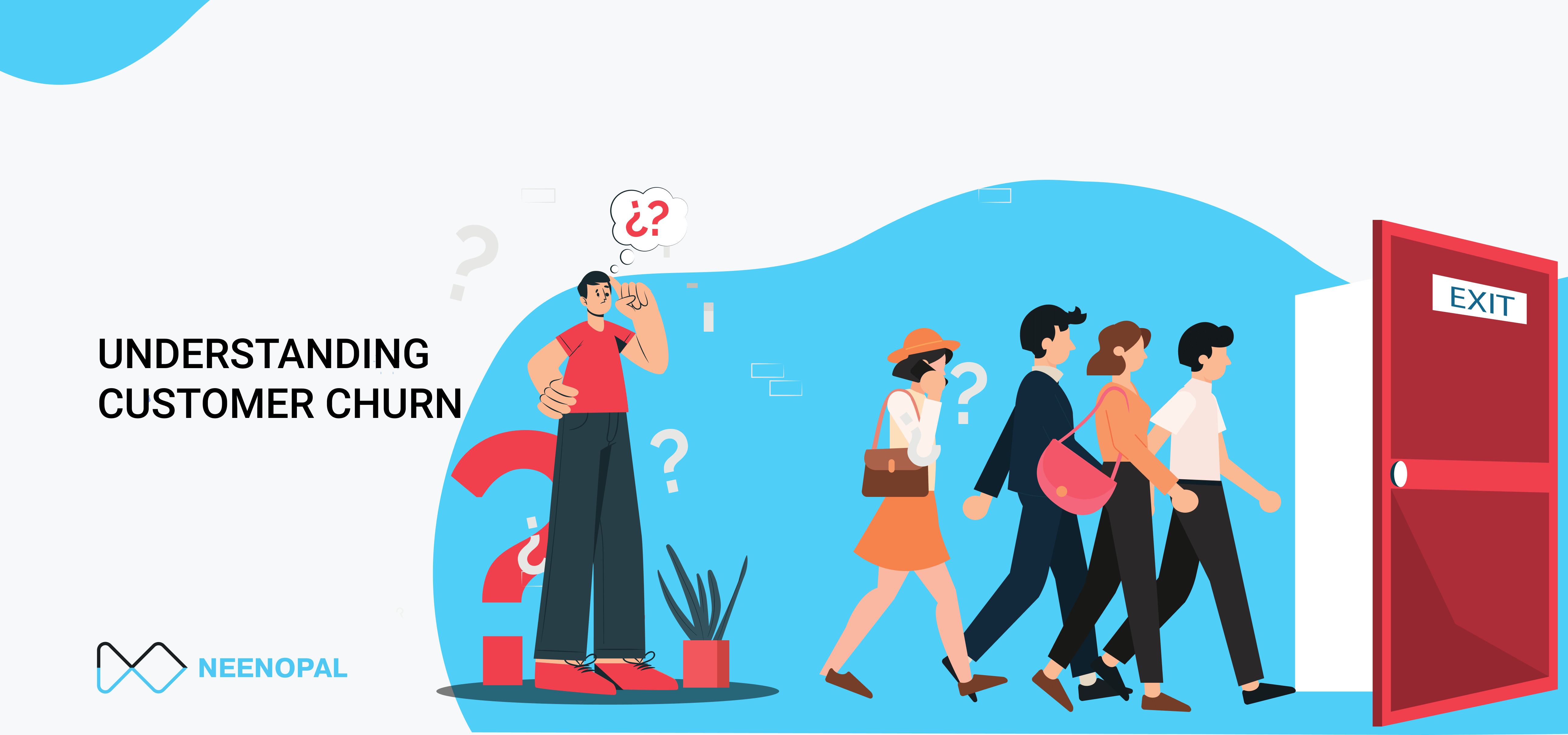
****

**Business Objective:**

Customer churn is a significant challenge for telecommunications companies, with annual churn rates often exceeding 10%. To address this issue, companies employ strategies to retain as many clients as possible. This project is a classification task, as the goal is to predict a binary outcome—whether a customer will churn or remain loyal. The objective is to model the probability of churn based on customer features.

**Data Set Details:**

Each row in the dataset represents a client of a company, with collected information including the type of plan, minutes of usage, and monthly charges

The dataset contains information on telecom customers, including state, location code, account duration, voicemail plan, voicemail messages, international plan, international call details, and usage metrics during the day, evening, and night. Other features encompass charges, call counts, and customer service interactions. The target variable, 'churn,' indicates whether the customer has left the company

This information is crucial for developing a predictive model to estimate the probability of churn based on customer characteristics

**Project Blueprint: Customer Churn Prediction**

**1. Data Collection:**

* Acquire a dataset containing customer information, including features such as account length, voice plan, international plan, etc.
* Ensure data integrity, check for missing values, and handle outliers if necessary.

**2. Exploratory Data Analysis (EDA):**

* Perform an in-depth exploration of the dataset.
* Visualize the data using various plots and charts.
  + Corelation plot: Explore correlations between variables.
  + Bar Plot: Visualize categorical variables.
  + Line Chart: Explore trends over time if applicable.
  + Histogram, Distribution Plots: Understand the distribution of numerical features.
  + Top and Bottom Analysis: Identify top and bottom-performing segments.

**3. Feature Engineering:**

* Clean the data and handle any missing values.
* Drop irrelevant columns that do not contribute to the model.
* Identify and select the most important features for model building.

**4. Model Building:**

* Implement various classification models:
  + Logistic Regression
  + Random Forest
  + Gradient Boosting
  + Support Vector Machines (SVM)
  + K-Nearest Neighbors (KNN)
  + Naive Bayes, etc.

**5. Model Evaluation:**

* Split the dataset into training and testing sets.
* Train each model using the training set.
* Evaluate models using metrics such as:
  + Confusion Matrix
  + Classification Report
  + Accuracy
  + Precision
  + ROC Curve (Receiver Operating Characteristic Curve)
  + AUC (Area Under the ROC Curve)
  + Recall

**7. Model Deployment:**

* Choose a deployment method (e.g., Streamlit, Flask).
* Deploy the chosen model for classification.
* Create a user-friendly interface for users to input customer information and get churn predictions.