## Прогнозування Відтоку Клієнтів для Телекомунікаційної компанії

### 1. Вступ

Нарешті ви влаштувались на першу роботу в IT на позицію Data Scientist. Ваш керівник вирішив дати вам невеличкий проєкт, щоб ви ознайомились з [базою даних клієнтів, видами сервісів, які надає компанія і технічними можливостями компанії](https://drive.google.com/file/d/121xd_NVxBIiDevkqnNUDpDdwj8jcR41D/view?pli=1).   
Цей документ описує технічне завдання проєкту, метою якого є розробка прогностичної моделі для ідентифікації ймовірності припинення клієнтами користування телекомунікаційними послугами на основі історичних даних про клієнтів. Проєкт передбачає використання аналізу даних, перед обробки даних, машинного навчання, оцінки моделей, а також пакування моделі у контейнер для забезпечення відтворюваності.

### 2. Опис Даних

Надається набір даних, який включає інформацію про клієнтів телекомунікаційної компанії: демографічні характеристики, історію використання послуг, тарифні плани, дані про відток(Churn).

### 3. Аналіз Даних (EDA)

* Використання бібліотек: Pandas, Matplotlib, Seaborn.
* Основні завдання: Вивчення розподілів, виявлення відсутніх значень, аналіз кореляцій.

### 4. Попередня Обробка Даних

* Обробка відсутніх значень: Заповнення або видалення.
* Кодування категоріальних змінних: One-Hot Encoding або Label Encoding.
* Нормалізація ознак: Стандартизація.

### 5. Розробка Моделі

* Вибір Алгоритму (ви можете використовувати будь-які алгоритми навчання які ви проходили на курсі)
* Навчання Моделі: Крос-валідація для параметрів.
* Оцінювання Моделі: Accuracy, Recall, Precision, F1 score

### 6. Інтеграція та Виведення Результатів

* Введення Даних Клієнта: Інтерфейс або процедура для даних нового клієнта.
* Опрацювання Моделлю: Аналіз даних і вирахування ймовірності відтоку.
* Вивід Результатів: "Клієнт має високу/низьку ймовірність відтоку". Візуалізація результатів.

### 7. Вимоги до проєкту

* Опис аналізу та обробки даних.
* Деталі моделі і обрані параметри.
* Аналіз результатів, метрик.
* Опис інтеграції та виведення результатів.
* Процес контейнеризації та інструкції користування.

### 8. Критерії Оцінювання

* Якість обробки даних і обґрунтування методів.
* Вибір параметрів моделі.
* Глибина аналізу результатів.
* Функціональність інтерфейсу даних і виведення.
* Успішна контейнеризація і документація.

### 9. Створення Dockerfile

* Розробіть Dockerfile для створення образу Docker, який дозволить розміщувати та запускати ваш проєкт в контейнеризованому середовищі. Dockerfile має включати всі необхідні інструкції для створення образу, включаючи вибір базового образу, копіювання вихідного коду програми до контейнера, встановлення необхідних залежностей та визначення команди для запуску програми.
* Інтегруйте інструмент Docker Compose для спрощення процесу розгортання та управління нашим проєктом у середовищі Docker. Створіть файл docker-compose.yml, який описує послуги, мережі та томи, необхідні для проєкту. Файл повинен дозволяти запускати весь проєкт за допомогою однієї команди docker-compose up, автоматизуючи створення та запуск необхідних Docker контейнерів.