1. **Tình huống:**
2. **Tình huống 1**

Bạn X đang sử dụng excel và dùng pandas để đọc dữ liệu, trong file excel hiện tại không có tên cột nhưng bạn X cần trao đổi với các bạn khác trong team mà không biết sử dụng code vì thế bạn X cần trao đổi dựa trên tên cột được xuất hiện mặc định trên excel như: A, B, C, .. , Z, AA

Nhưng khi sử dụng code bạn X không biết được cột đó ở trong Excel như thế nào mà chỉ biết là index của cột là 1,2,3,4. Vì thế để việc trao đổi hiệu quả bạn X cần biết cột tương ứng từ pandas sang Excel: 1 -> A, B -> 2 …

Làm cách nào để bạn X có thể biết được tên một cột bất kì trong excel: như XA, ZA, AAA, `` **Bạn có thể giúp bạn X tìm được tên cột dựa trên index không?**

1. **Tình huống 2**

Nhà bạn Y làm trong ngành in ấn biển quảng cáo và banner, hàng ngày bạn Y cần phải in hàng nghìn chữ để dán lên biển quảng cáo với các thông điệp khác nhau. Tuy nhiên trong quá trình làm việc bạn Y hay sử dụng điện thoại để chơi game dẫn đến việc in sai các thông điệp. Thông thường bạn Y sẽ in lại cả tiếng hoặc từ hững sau đó bạn nhận ra việc in lại toàn bộ từ và tiếng là tốn kém và lãng phí vì vậy sau khi in ra toàn bộ các thông điệp bạn ấy có thể in riêng lẻ các chữ để thay thế sau đó cuối ngày bạn ấy cần tính toán lại tổng số chữ đã in thừa / thiếu riêng lẻ để bổ sung. Ví dụ:

Biển qung céoi -> Biển quảng cáo: Cần in thêm 1 chữ, thay thế 1 chữ, và thừa 1 chữ -> Tổng số chữ in riêng lẻ và thừa là 3.

**Bạn có thể giúp bạn Y tính nhanh được số chữ in riêng lẻ và thừa không?**

1. **Phân tích**

**Nhiệm vụ:**

1. Sử dụng bộ dữ liệu được cung cấp để phân tích và phát triển mô hình học máy nhằm xác định khả năng default (target=1) của phương án vay tại tổ chức tín dụng.
2. Gửi lại source code python & slide mô tả cách triển khai xử lý dữ liệu -> Phát triển mô hình -> Đánh giá hiệu năng
3. Giải tích ý nghĩa các biến trong mô hình final; tác động của các biến đến mô hình

**Bộ dữ liệu bao gồm:**

* application\_{train|test}.csv: Bảng chính, chứa thông tin các phương án vay của Khách hàng tại tổ chức tín dụng. (train có target; test không có target)
* bureau.csv: Thông tin CIC về các khoản vay tín dụng của Khách hàng tại các tổ chức tín dụng khác.
* bureau\_balance.csv: Thông tin CIC về dư nợ của Khách hàng tại các tổ chức tín dụng khác.
* POS\_CASH\_balance.csv: Sao kê hàng tháng số dư khoản vay và tài khoản thanh toán của Khách hàng.
* credit\_card\_balance.csv: Sao kê hàng tháng thẻ tín dụng của Khách hàng đã có tại tổ chức tính dụng.
* previous\_application.csv: Lịch sử tất cả các phương án vay trước đây của khách hàng.
* installments\_payments.csv: Lịch sử trả nợ các khoản tín dụng đã giải ngân trước đó.
* columns\_description.csv: mô tả các cột trong các file dữ liệu trên.

Link Dataset: [Credit Default (kaggle.com)](https://www.kaggle.com/datasets/congng/credit-default)