最終課題：債務不履行の予測

以下のようなデータから、クレジットカードの返済延滞が起こるかどうかを予測してください。

* **train.csv**: 教師データです。説明変数と正解が与えられています。これらからモデルを作成して、test.csvに対する予測を作成します。
* **test.csv**: これらのデータにたいする予測を作成してサブミットしてください。
* **sample\_submission.csv**: サブミッション用ファイルのサンプルです。この形式のファイルをサブミットしてください

# データの概要

次月の返済延滞（default payment next month）（Yes = 1, No = 0）を予測してください。以下の 23 の説明変数があります。

'LIMIT\_BAL': 与信額（NT ドル）：個人の消費者信用と家族の（補足的な）信用の両方を含む。

'SEX': 性別（1＝男性、2＝女性）

'EDUCATION': 学歴（1=大学院、2=大学、3=高校、4=その他）

'MARRIAGE': 婚姻状況（1＝既婚者、2＝独身、3＝その他）

'AGE': 年齢（年）

'PAY\_0'～'PAY\_6'：過去の支払い履歴。過去の月々の支払い履歴（2005年4月～9月）を以下のように追跡しました。

'PAY\_0'＝2005年9月の返済状況、'PAY\_1'＝2005年8月の返済状況、...... 'PAY\_6'＝2005年4月の返済状況。

　これらの返済状況の測定尺度は、-1＝正当な支払い、1＝1ヶ月間の遅延、2＝2ヶ月間の遅延、......、8＝8ヶ月間の遅延、9＝9ヶ月以上の遅延となっています。

'BILL\_AMT1'～'BILL\_AMT6'：手形明細書の金額（NTドル）。'BILL\_AMT1'＝2005年9月の手形明細の金額、'BILL\_AMT2'＝2005年8月の手形明細の金額、...... 'BILL\_AMT6'＝2005年4月の手形明細の金額。

'PAY\_AMT1'～'PAY\_AMT6'：過去の支払額（NTドル）。'PAY\_AMT1'＝2005年9月に支払った金額、'PAY\_AMT2'＝2005年8月に支払った金額; ...... 'PAY\_AMT6'＝2005年4月の手形明細の金額。

'default payment next month': 次月に返済延滞があったか否か（0,1）[target]。

# URL

以下のURLから予測を入れたファイルをsubmitしてください。

<https://www.kaggle.com/t/01b41fdc2e3b4787930fa7243c9deaad>

# 締め切り

1月15日, 24:00締め切りとします。この日までに、上記URLに予測ファイルをサブミットしたうえで、manabaに予測モデルを作ったipynbファイルを提出してください。

* もし事情があって遅れそうなときは、早めに連絡してください。

# 評価

コンペティションの順位とipynbファイルの内容の両方を考慮します。

これまでの授業内の課題提出とこの最終課題で成績はつけます。

## 連絡先

質問や気になることがありましたら、福田（haruaki.fukuda@r.hit-u.ac.jp）まで連絡してください。