# Machine Learning - project

**Abstract**

Credit card default happens when you've become severely delinquent on your credit card payment.

It's a serious credit card status that not only affects your standing with that credit card issuer, but also your credit standing in general and your ability to get approved for credit cards, loans, and other credit-based services.

When you accept a credit card, you agree to certain terms, e.g. you agree to make your minimum payment by the due date listed on your credit card statement.

If you miss the minimum credit card payment six months in a row, your credit card will be in default; your credit card issuer will likely close your account and report the default to the credit bureaus.

By the time your credit card defaults, you've likely accumulated hundreds of dollars in fees and interest charges.

Unfortunately, your options for clearing up the credit card default may be limited because of the number of payments you've missed on your account.

For this reason, assuming truthful the given data, we show the procedure used to create prevision algorithm aiming to foresee the default payment's client.

## 1 Introduction

This dataset is available on Kaggle under the name *Default of Credit Card Clients Dataset* and it contains information on default payments, demographic factors, credit data, history of payment and bill statements of credit card clients in Taiwan from April 2005 to September 2005.

The following variables and their characteristics are:

- ID: ID of each client
- LIMIT_BAL: Amount of given credit in NT dollars
- SEX: Gender (1=male, 2=female)
- EDUCATION: (1=graduate school, 2=university, 3=high school, 4=others, 0,5,6=unknown)
- MARRIAGE: Marital status (1=married, 2=single, 3=divorced, 0=others)
- AGE: Age in years
- PAY_0: Repayment status in September, 2005 (-2=no consumption, 0=use of revolving credit, -1=pay duly, 1=payment delay for one month, 2=payment delay for two months, ... 8=payment delay for eight months, 9=payment delay for nine months and above)

- PAY_2: Repayment status in August, 2005 (scale same as above)
- PAY_3: Repayment status in July, 2005 (scale same as above)
- PAY_4: Repayment status in June, 2005 (scale same as above)
- PAY_5: Repayment status in May, 2005 (scale same as above)
- PAY_6: Repayment status in April, 2005 (scale same as above)
- BILL_AMT1: Amount of bill statement in September, 2005 (NT dollar)
- BILL_AMT2: Amount of bill statement in August, 2005 (NT dollar)
- BILL_AMT3: Amount of bill statement in July, 2005 (NT dollar)
- BILL_AMT4: Amount of bill statement in June, 2005 (NT dollar)
- BILL_AMT5: Amount of bill statement in May, 2005 (NT dollar)
- BILL_AMT6: Amount of bill statement in April, 2005 (NT dollar)
- PAY_AMT1: Amount of previous payment in September, 2005 (NT dollar)
- PAY_AMT2: Amount of previous payment in

August, 2005 (NT dollar)

- PAY_AMT3: Amount of previous payment in July, 2005 (NT dollar)
- PAY_AMT4: Amount of previous payment in June, 2005 (NT dollar)
- PAY_AMT5: Amount of previous payment in

May, 2005 (NT dollar)

- PAY_AMT6: Amount of previous payment in April, 2005 (NT dollar)
- default.payment.next.month: Default payment (1=yes, 0=no)