

The Dataset

The dataset for ML model building has the following columns:

Data type: int64	Variables	Info	Key
Target	credit	Recoded from LIMIT_BAL	Both the individual consumer credit and his/her family (supplementary) credit.
Features	default	Recoded from Default	0 = not default, 1=default
	AGE	Age (year)	
	MARRIAGE	Marriage status	1 = married; 2 = single; 3 = divorce; 0=others
	EDUCATION_ graduate school	Dummy variable from categorical EDUCATION	0, 1
	PAY_2	History of past payment, (from April to September, 2005) The repayment status for one in August, 2005	1: Paid in full; 0: The use of revolving credit; 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.
	BILL_AMT5	amount of bill statement for one of the six months (NT dollar): May, 2005	
	PAY_AMT6	Amount of previous payment (NT dollar) for one of the six months: April, 2005	