Give Me Some Credit Data set

July 1, 2016

1 Background

'GiveMeSomeCredit' training dataset is hosted on Kaggle which contains approximately 150,000 observations, with 11 related variables. There is also a test set which includes 101,503 instances. The goal is to determine whether or not a loan should be granted.

2 Attributes

ydata: (binary) Whether or not to grant the loan for this user. The target variable of training data set. Value: 0,1.

Revolving Utilization Of Unsecured Lines: (numerical) Total balance on credit cards and personal lines of credit except real estate and no installment debt like car loans divided by the sum of credit limits. Range from: 0.0-50708.0. **Lower is better**

Age: (numerical) Age of borrower in years Range from: 0-109.

Number Of Time 30-59 Days Past Due Not Worse: (numerical) Number of times borrower has been 30-59 days past due but no worse in the last 2 years. Range from: 0-98.**Lower is better**

DebtRatio: (numerical) Monthly debt payments, alimony,living costs divided by monthy gross income. Range from: 0.0-329664.0. **Lower is better**

MonthlyIncome: (numerical) Monthly income. Range from: 0.0-3008750.0. **Higher is better**

NumberOfOpenCreditLinesAndLoans: (numerical) Number of Open loans (installment like car loan or mortgage) and Lines of credit (e.g. credit cards) Range from: 0-58. **Lower is better**

NumberOfTimes90DaysLate: (numerical) Number of times borrower has been 90 days or more past due. Range from: 0-98. **Lower is better**

NumberRealEstateLoansOrLines: (numerical) Number of mortgage and real estate loans including home equity lines of credit. Range from: 0-54. **Lower is better**

NumberOfTime60-89DaysPastDueNotWorse: (numerical) Number of times borrower has been 60-89 days past due but no worse in the last 2 years. Range from: 0-98. **Lower is better**

NumberOfDependents: (numerical) Number of dependents in family excluding themselves (spouse, children etc.). Range from: 0-20. **Lower is better**