PREDICTING IMBALANCED CLASSIFICATION

Front Matter:

* Misclassification errors on the minority class are more important than other types of prediction errors for some imbalanced classification tasks.
* One example is the problem of classifying bank customers as to whether they should receive a loan or not. Giving a loan to a bad customer marked as a good customer results in a greater cost to the bank than denying a loan to a good customer marked as a bad customer.

Overview

1. Credit Dataset
2. Explore Dataset
3. Model test and Result
4. Evaluate Models
5. Make prediction on new data

Dataset:

The credit dataset describes financial and banking details for customers and the task is to determine whether the customer is good or bad. The assumption is that the task involves predicting whether a customer will pay back a loan or credit.

The dataset includes 1,000 examples and 20 input variables, 7 of which are numerical (integer) and 13 are categorical.

* Status of existing checking account
* Duration in month
* Credit history
* Purpose
* Credit amount
* Savings account
* Present employment since
* Installment rate in percentage of disposable income
* Personal status and sex
* Other debtors
* Present residence since
* Property
* Age in years
* Other installment plans
* Housing
* Number of existing credits at this bank
* Job
* Number of dependents
* Telephone
* Foreign worker