Finance Risk Model (Logistic Regression Model)

Qi Dong

April 14, 2019

(Note: This project is still under development, which means there could have bugs, problems, or even function declaration-and-deployment inconsistency.)

1 Introduction

2 PYTHON Files

For the general usage, please refer to "executer.py" for automatic LR modeling. It includes the complete LR modeling execution process, including feature inspection and selection, feature elimination based on Lasso method or recursive feature elimination (RFE) or others, and LR modeling with L2 regularization (my personal preference). Of course there are different ways of LR modeling, but such a routine is sufficient for finance risk modeling in a majority of situations.

Following provides a brief explanation of each individual files (3 included and others are ignored):

2.1 "feature_inspection.py"

"feature_inspection.py" is mainly used for a). basic statistic exploration of each features (such as ks, auc, missing rate, zero rate, etc.), b). missing data padding, and c). rudimentary feature encoding if necessary (please note: more sophisticated encoding techniques might be invoked in particular cases).

2.2 "binning.py"

"binning.py" divides features into bins, explores more statistical features such as IV, PSI, bad rate of each bin, etc., and compute woe of each bin of features.

2.3 "model_inspection.py"

"model_test.py" serves for the model evaluation including ROC generation, KS and AUC computation, and lift chart plotting.