Discussion of “Corporate Default Forecasting with Machine Learning” by Mirko Moscatelli

Isaiah Hull

Sveriges Riksbank

October 20, 2019

Research Question

Can we improve corporate default forecasting by using machine learning models, rather than standard statistical models?

Overview

Models

Standard Models

- Linear Discriminant Analysis
- Logistic Regression

Machine Learning Models

- Penalized Logistic Regression
- Random Forest
- Gradient Boosted Trees

Overview

Results

Discriminatory Power: ML performance better under two conditions:

- High number of observations.
- Low number of features.

Credit Allocation: Positive impact on amount of credit extended.

- Reduced credit losses for lenders.
- Limits lending to smaller pool of borrowers.

Comments

Comments

- Well-executed and carefully explained.
- Explains benefits of ML for important and relevant empirical problem.
- Avoids overstating the power of ML by performing nuanced comparison.

Comment #1

Variable Selection Criteria

1. *Using univariate logit regressions for the probability of default, those variables with AuROC lower than fifty-five per cent were dropped*

2. *Using the Kolmogorov-Smirnov test, those variables with insignificant differences in the distributions between the default and non-default groups were dropped*

3. *From the list satisfying 1) and 2), only less correlated variables were retained (linear correlation < 0.7)*

Comment #1

Variable Selection Criteria

Figure 1: Decision Tree

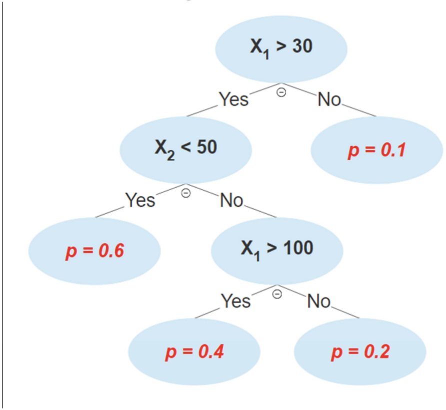

Figure taken from Moscatelli (2019).

Comment #1

Variable Selection Criteria

- Nonlinearities <u>between</u> features may also be important.
- Using univariate selection criteria could be too restrictive.

Comment #2

Feature Engineering

- Albanesi and Vamossy (2019) find substantial gains from using deep neural networks to predict consumer default.
- Random forests allow for a more limited form of nonlinearity.
  - Conditioning thresholds.
- May need feature engineering to realize full gains of ML.
  - Kaggle: Aggregation, ratios, sums, differences.
  - Deep Feature Synthesis (Kanter and Veeramachaneni, 2015).
  - Python: Featuretools.

Comment #3

Interpretation of Results

- Adding more features weakens the advantage of ML over standard models.
- This is contrary to expectations, given outcomes of ML competitions and findings in ML literature.
- What explains this result?

Comment #3

Interpretation of Results

- High number of observations and relatively low number of features.
  - Feature engineering.
- Importance of credit behavioral indicators.
  - Consider credit behavioral indicators in isolation.
- Nonlinearity in small number of features.
  - Explore other ML models.