Predicting bank insolvencies using Machine Learning techniques

2017 EBA Policy Research Workshop

London, 28 November 2017

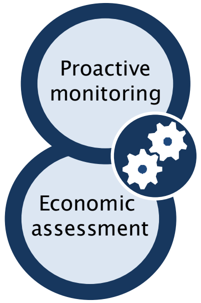
BANK OF GREECE
Anastasios Petropoulos
Vasilis Siakoulis
Evangelos Stavroulakis
Nikolaos E. Vlachogiannakis

The views expressed in this paper are those of the authors and not necessarily those of Bank of Greece

Predicting bank insolvencies

In a nutshell

Novel Rating System
for financial institutions



Modeling technique

- Random Forests

Main Drivers

- Profitability
- Capital
- Asset Quality

Implementation

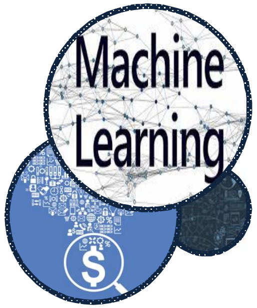
- Development sample US bank data (source: FDIC)
- Testing sample European bank data (source: SNL)

Predicting bank insolvencies
Motivation

Supervisory Authorities Priorities

- Protection of depositors interest
- Verification of Financial Institutions ability to survive under normal and stressed conditions

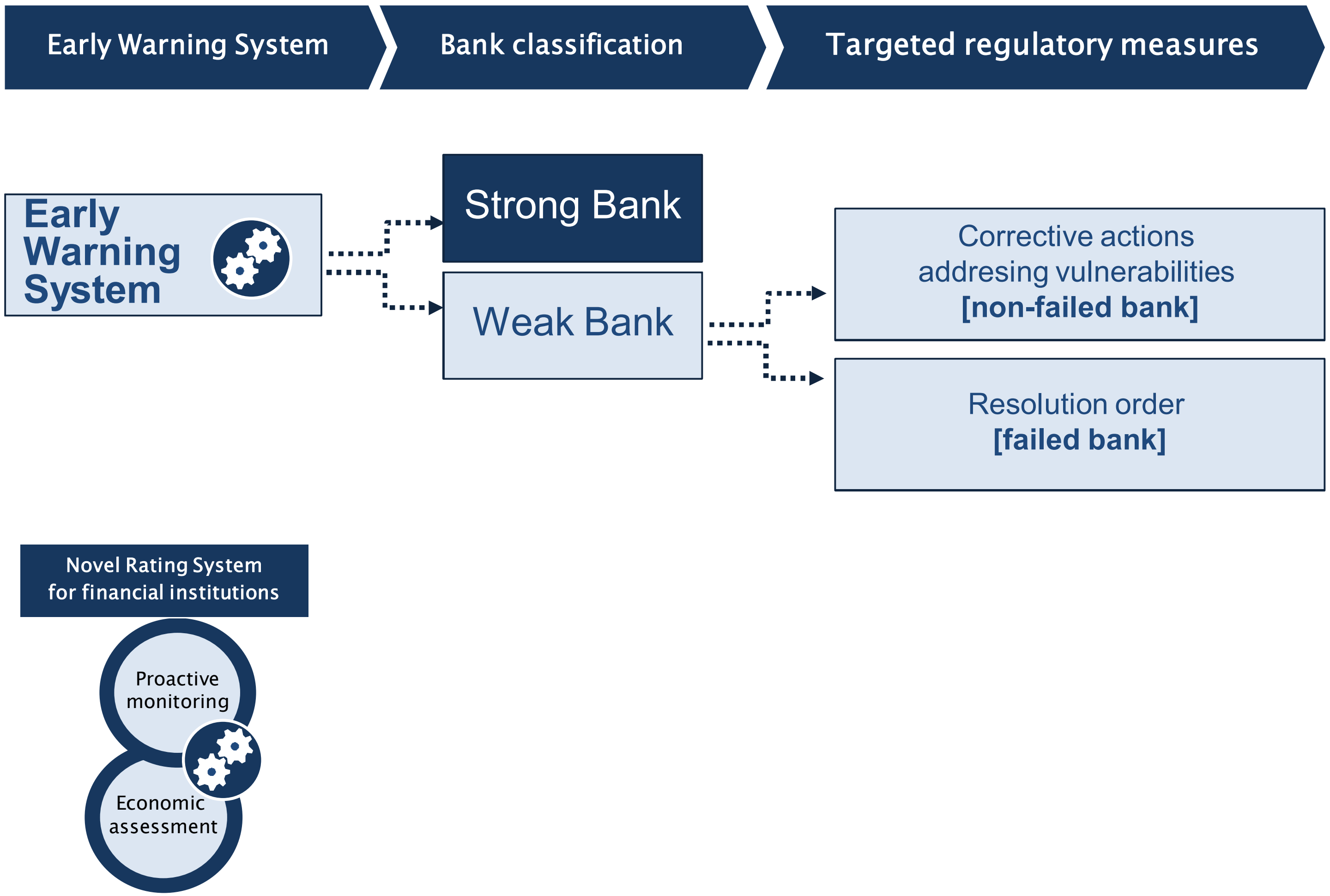
Predicting bank insolvencies
Machine learning techniques (ML)



- Machine learning techniques (ML) are related to the capacity of the computers to “learn” without being explicitly programmed.
- ML explore the study and construction of algorithms that can learn from, and make predictions on data for which designing and programming explicit algorithms with good performance is difficult or infeasible.
- Many use cases of ML are recently favored by the technological advances, the availability of financial sector data and infrastructure.
- They are being rapidly adopted for a range of applications in the financial services industry, as they are:
  - more efficient in processing of information,
  - able to identify new determinants,
  - capable of unveiling unexpected forms of dependencies among variables.
- Therefore, supervisory authorities should keep up with the current developments so as not only to uncover "hidden insights” of modelling relationships, but also to be able to benchmark/challenge the results of the supervised entities.

Predicting bank insolvencies

The ultimate goal



Predicting bank insolvencies

The story so far

4,069 failed banks

Banks in the US that failed or received financial assistance from FDIC, in the 1943 – 2014 period.

3,483 failed banks

Banks in the US that failed or were assisted by the Central Bank, in the 1980 – 2014 period.

Basel III

Due to the 2008 financial crisis the Basel II requirements were further strengthened.

The following elements were also introduced:

i. Counter-cyclical capital buffer

ii. Leverage ratio (non-risk weighted)

iii. Liquidity requirement (i.e. LCR and NSFR)

It is clear that premeptive identification of insolvent banks was not so effective and supervisory authorities should further strengthen the monitoring process of the banking system

- The compliance with an even more extended set of minimum regulatory standards and/or the monitoring of the banks risk indicator, should not be assessed on a standalone basis.
- All risk drivers and relevant information should be combined in a single measure.
- This is a difficult task due to the big bulk of information that is currently collected

Supervisory authorities should utilize robust aggregation methodologies, resulting in an efficient calculation of a survival probability for each financial institution & its classification into different risk classes.

Predicting bank insolvencies

Current Approaches and Practices

Statistical Methods

Methods used in the last decades to aggregate bank specific information in a single figure

- Discriminant analysis
- Logit/ Probit regressions
- Advanced Machine Learning
  - Support Vector Machines
  - Conditional Inference Trees
  - Neural Networks

Random Forests

Popular method for modeling classification problems in recent years

- Other novel modelling approaches such as Random Forest have not been employed up to now in assessing bank failures.
- Random Forests is a useful framework for analysis of big datasets and handling a large number of input variables without any correlation restrictions.
- Random Forests are also efficient in modeling outliers due to the random subspaces process and their ability to recognize nonlinear relationships.

However

No academic study exists that thoroughly assess simultaneously all those methodologies on a common dataset, in order to determine in a concrete way their relative forecasting performance

Predicting bank insolvencies

Our Approach

Modeling Approaches

- Logistic Regression
- Linear Discriminant Analysis
- Random Forests
- Support Vector Machines
- Neural Networks
- Random Forests of Conditional Inference Trees

Performance Metrics

- Model performance on imbalanced samples
  - In-sample
  - Out-of-sample
  - Out-of-time

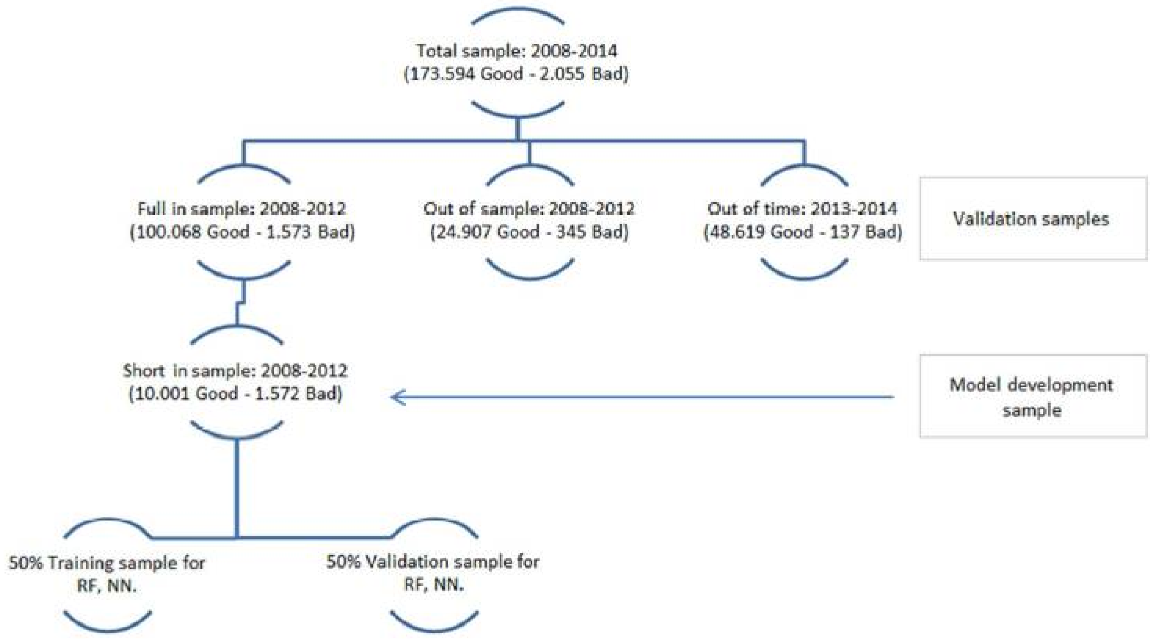
Dataset

- FDIC data:
  - 2008-2014: 175,649 records
- Bank specific information (drivers)
  - 40+ CAMELS variables
  - 660+ covariates analyzed and examined

- Recent empirical evidence suggests that the financial condition of banks is the key driver in distinguishing their performance during stress periods.
- Supervisory authorities are mainly interested in identifying the bank-specific weakness that may drive banks to insolvency, so that they can address them via the specification of targeted remediation actions.

Data
Samples used

- FDIC sourced data related to all US banks.
- The definition of a default event includes all bank failures and assistance transactions of all FDIC-insured institutions.
- The total sample covered the 2008-2014 period, including 173,594 Good and 2,055 Bad (175,649 total records included in the analysis)



Potential drivers

CAMELS

C — Capital adequacy

- Equity capital to assets
- Core capital (leverage) ratio
- Tier 1 risk-based capital ratio
- Total risk-based capital ratio
- Common equity tier 1 capital ratio

A — Asset quality

- Loan and lease loss provision to assets
- Net charge-offs to loans
- Credit loss provision to net charge-offs
- Assets per employee ($millions)
- Earning assets to total assets ratio
- Loss allowance to loans
- Loan loss allowance to noncurrent loans
- Noncurrent assets plus other real estate owned to assets
- Noncurrent loans to loans
- Average total assets
- Average earning assets
- Average equity
- Average total loans
- Net loans and leases
- Loan loss allowance
- Restructured Loans & leases
- Assets past due 30-89 days
- Restructuring ratio
- Provisions to loans
- Provision to assets

M — Management capability

- Noninterest income to average assets
- Noninterest expense to average assets
- Net operating income to assets
- Earnings coverage of net charge-offs
- Efficiency ratio
- Cash dividends to net income

E — Earnings

- Yield on earning assets
- Cost of funding earning assets
- Net interest margin
- Return on assets
- Pretax return on assets
- Return on Equity
- Retained earnings to average equity

L — Liquidity

- Net loans and leases to total assets
- Net loans and leases to deposits
- Net loans and leases to core deposits
- Total domestic deposits to total assets
- Volatile Liabilities

S — Sensitivity to Market Risk

- Asset Fair Value

Potential drivers

Variable reduction process

660

- **Correlation with dependent variable:** From 660 initial predictors the variables exhibiting the highest in-sample correlation with the dependent variable kept, based on a threshold of at least +/-10% correlation.

104

- **Cross correlation analysis:** Variables with pair-wise correlation over 60% in absolute value with another variable excluded. Among pairs of collinear variables the one exhibiting higher predictive ability with the dependent variable remained.

59

- **LASSO process:** Assess simultaneously the explanatory power of the remaining variables. Number of regressors to be tested decreased to 59.

23

- **Random Forest Importance:** Rank variables based on Random Forest variance importance plot. Number of regressors to be tested decreased to 23.

Performance results

What is the best statistical approach

Table 1: Short in-sample performance metrics

	Logit	LDA	RF	SVM	NN	CRF
AUROC	0,980	0,973	0,989	0,981	0,984	0,991
G-mean	0,898	0,884	0,921	0,898	0,923	0,914
LR-	0,183	0,209	0,139	0,184	0,137	0,156
DP	3,116	2,971	3,255	3,181	3,356	3,312
BA	0,902	0,889	0,923	0,902	0,925	0,916
Youden	0,804	0,778	0,846	0,804	0,851	0,833
WBA1	0,943	0,936	0,953	0,944	0,955	0,951
WBA2	0,861	0,842	0,893	0,860	0,895	0,881

Table 3: Out-of time performance metrics

	Logit	LDA	RF	SVM	NN	CRF
AUROC	0,990	0,974	0,976	0,993	0,990	0,965
G-mean	0,741	0,824	0,862	0,819	0,862	0,838
LR-	0,452	0,321	0,255	0,329	0,255	0,296
DP	3,684	3,590	3,793	3,804	3,722	3,668
BA	0,774	0,839	0,871	0,835	0,871	0,851
Youden	0,548	0,677	0,743	0,670	0,742	0,702
WBA1	0,886	0,918	0,934	0,916	0,934	0,924
WBA2	0,662	0,759	0,809	0,754	0,809	0,778

Table 2: Out-of sample performance metrics

	Logit	LDA	RF	SVM	NN	CRF
AUROC	0,990	0,983	0,990	0,992	0,980	0,989
G-mean	0,919	0,905	0,934	0,916	0,922	0,907
LR-	0,144	0,169	0,113	0,150	0,130	0,165
DP	3,239	3,099	3,352	3,268	3,051	3,147
BA	0,921	0,908	0,935	0,919	0,923	0,910
Youden	0,842	0,816	0,871	0,837	0,847	0,821
WBA1	0,952	0,945	0,959	0,952	0,948	0,947
WBA2	0,890	0,871	0,912	0,886	0,898	0,874

Table 4: Full in-sample performance metrics

	Logit	LDA	RF	SVM	NN	CRF
AUROC	0,980	0,973	0,998	0,981	0,981	0,990
G-mean	0,898	0,884	0,992	0,897	0,926	0,914
LR-	0,184	0,209	0,000	0,185	0,125	0,153
DP	3,079	2,960	Inf	3,115	3,124	3,202
BA	0,901	0,889	0,992	0,901	0,927	0,916
Youden	0,803	0,777	0,984	0,802	0,854	0,832
WBA1	0,942	0,935	0,988	0,943	0,951	0,950
WBA2	0,860	0,842	0,996	0,859	0,903	0,883

- The performance results in all samples, indicate that the **Random Forest** rating system exhibits higher discriminatory power compared to all the considered benchmark models when taking into account the skewness of the data.
- More importantly, the obtained performance is more stable and more consistent across all test samples, resulting in lower performance variability.
- Another interesting finding stemming from our results is that Neural Networks perform relatively well in the “in-sample” and “out-of-time” samples.

Performance results

What is the main driver of default?

Main Default Indicators

- There is a big debate in the literature regarding the superiority of certain indicators in predicting bank failures:
  - Leverage Ratio (unweighted) vs.
  - Capital Adequacy Ratio (risk weighted)

- We include separately in the two drivers in our specification and compare model performance.
- Comparison is based on:
  - AUROC performance metric for Logit, LDA, SVM, NN and
  - MSE% variable importance plot for RF and CRF.

Dominance of Capital Adequacy Ratio vs Leverage Ratio per model.

	CAR	LEV
Logit		✓
LDA		✓
RF	✓	
SVM	✓	
NN	✓	
CRF	✓	

Our analysis implies that the importance of the one indicator relative to the other is purely model driven.

Performance results

What is the main driver of default?

We rank the predictors used across all models developed based on their marginal contribution (contribution to AUROC metric).

- Specifically, we excluded each variable, in turn, from each model and we measured the loss in AUROC for each specification.
- We ranked first the variables that led to the largest loss in AUROC metric.
- Our results indicate that metrics related to capital and earnings constitute the factors with the highest marginal contribution in predicting bank failures.

Covariate importance ranking per model

(1: Highest importance, 11: Lowest importance)

	Logit	LDA	RF	SVM	NN	CRF	Average Score	
log(equity)(-4)%	4	3	11	3	7	6	5,7	Equity (log of Equity previous 4Q)
d(LEV)(-4)	3	5	8	5	6	8	5,8	Leverage (change during last 4Q)
LOSS_LOAN_DFS	7	6	10	1	1	10	5,8	Loss allowance to loans (difference from sector average)
d(NCASS_ORE)(-4)	9	10	2	8	4	4	6,2	Noncurrent assets plus other real estate owned to assets (change during last 4 Q)
d(ROA)(-4)	11	8	4	6	8	9	7,7	Return on Assets (change during last 4Q)
LEV	2	2	3	10	5	3	4,2	Leverage
NLOAN_CDEP	6	9	7	2	3	7	5,7	Net loans and leases to core deposits
NPL	8	7	5	9	11	5	7,5	Noncurrent loans to loans
LOSS_NPL	10	11	6	11	9	11	9,7	Loan loss allowance to noncurrent loans
ROE	5	4	9	4	10	1	5,5	Return on Equity
CFEA	1	1	1	7	2	2	2,3	Cost of funding earning assets

Logit	Logistic Regression
LRA	Linear Discriminant Analysis
RF	Random Forests
SVM	Support Vector Machines
NN	Neural Networks
CRF	Random Forest of Conditional Inference Trees

Our Model
Case study and Benchmark

- We applied the Random Forests specification in creating an Early Warning System of bank failures in Europe.
- This is a strong test for classification purposes as this region is characterized by significant disparity in financial institutions driven by country macroeconomic specificities.
- Specifically, we calculated Probability of Default for 173 European banks based on year end-2015 accounting and regulatory data (source: SNL).
  - **The analysis is indicative and the results are subject to data availability and quality.**

- In order to benchmark our results we mapped our PDs to rating classes based on lower bound PD thresholds described in 2016 Moody's rating methodology document.
- We evaluated the concordance of our ranking with the respective Moody's ranking by calculating Kendal's tau, Spearman's rho and the classical Fisher correlation coefficient.
  - Seeing that Moody's ratings take into account the sovereign rating of a bank's resident country, we adapted our ranking for sovereign rating in a similar way as described in Moody's respective document .

- Our credit rating scale has 67% Spearman's Rho, 59% correlation and 47% Kendal's Tau with the Moody's Rating system, thus, verifying the high positive concordance.

Our Model
Case study results

	High Risk Banks	Banks in Sample
AT	0	5
BA	0	2
BE	1	2
BG	1	2
CH	0	17
CY	0	2
CZ	0	1
DE	2	8
DK	0	15
ES	1	8
FI	0	2
FR	3	11
GB	0	10
GE	0	1
GR	3	5
HR	2	3
HU	2	2
IE	0	3
IT	7	15
LI	0	1
MD	1	1
MK	0	2
MT	0	3
NO	0	12
PL	0	11
PT	1	2
RO	1	2
RS	0	1
RU	5	5
SE	1	3
SK	0	4
TR	7	9
UA	3	3
Grand Total	**41**	**173**

- The table shows the number of High Risk banks by country. A bank is defined as High Risk when its Probability of Default, as calculated by the Random Forest specification, is larger than 25%.
- Focusing on Eurozone we notice that countries experiencing prolonged macroeconomic deterioration, which has eroded local banks' capital and increased non-performing exposures, show the highest relative number of "High Risk banks".
- On the other side, our results confirm that stronger Eurozone economies are accompanied by resilient banking systems.
- Finally, countries regaining competitiveness exhibit relative low levels of risky banks.
- Outside Eurozone, strong economies exhibit close to zero levels of "High Risk Banks".
- We finally remain cautious on the results in Eastern European countries and small countries for which our sample contains a limited number of banks.

Predicting bank insolvencies using Machine Learning

Our Contribution

- ✓ Extensive exploration of statistical techniques
- ✓ First empirical application of Random Forests
- ✓ Robust validation
- ✓ Performance measures
- ✓ Extended set of potential drivers

Thank you!