

Machine Learning in Economics
(458657)

replication paper

Early Warning System of Fiscal Stress

**comparing the traditional logistic regression approach versus a
random forest algorithm**

**Department of Economics
University of Bern**

Spring Semester 2022

submitted by Bela Koch

Contents

1	Introduction	3
2	Literature Review	3
3	Model Description	3
3.1	Performance Metrics	3
3.2	Logit Model with LASSO penalisation	3
3.3	Random Forest	3
4	Data Description	3
4.1	Dependent Variable	3
4.2	Explanatory Variables	3
5	Empirical results	3
5.1	Performance	4
5.2	Interpretability	4
5.2.1	Variable Importance	4
5.2.2	Shapley Values	4
5.2.3	Partial dependence plots	4
5.2.4	Accumulated local effects plots	4
6	Conclusion	4
7	References	5

1 Introduction

DEFINITION OF EWS. This paper aims to design an early warning system which signals increased risk of a fiscal stress event in the near future.

test ob zitierung funktioniert Jarmulska (2020).

2 Literature Review

3 Model Description

3.1 Performance Metrics

3.2 Logit Model with LASSO penalisation

Hastie et al. (2009)

$$\hat{\beta}^{lasso} = \underset{\beta}{\operatorname{argmin}} \sum_{i=1}^N (y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j)^2 \quad \text{subject to} \quad \sum_{j=1}^p |\beta_j| \leq t \quad (1)$$

Lagrangian form

$$\hat{\beta}^{lasso} = \underset{\beta}{\operatorname{argmin}} \left\{ \frac{1}{2} \sum_{i=1}^N (y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j)^2 + \lambda \sum_{j=1}^p |\beta_j| \right\} \quad (2)$$

3.3 Random Forest

Gini index

$$g(w) = \sum_{k \neq j} p_{wk} p_{wj} = \sum_k p_{wk} (1 - p_{wk}) \quad (3)$$

4 Data Description

4.1 Dependent Variable

definition of a fiscal stress event

empirical/historical data about fiscal stress events

4.2 Explanatory Variables

5 Empirical results

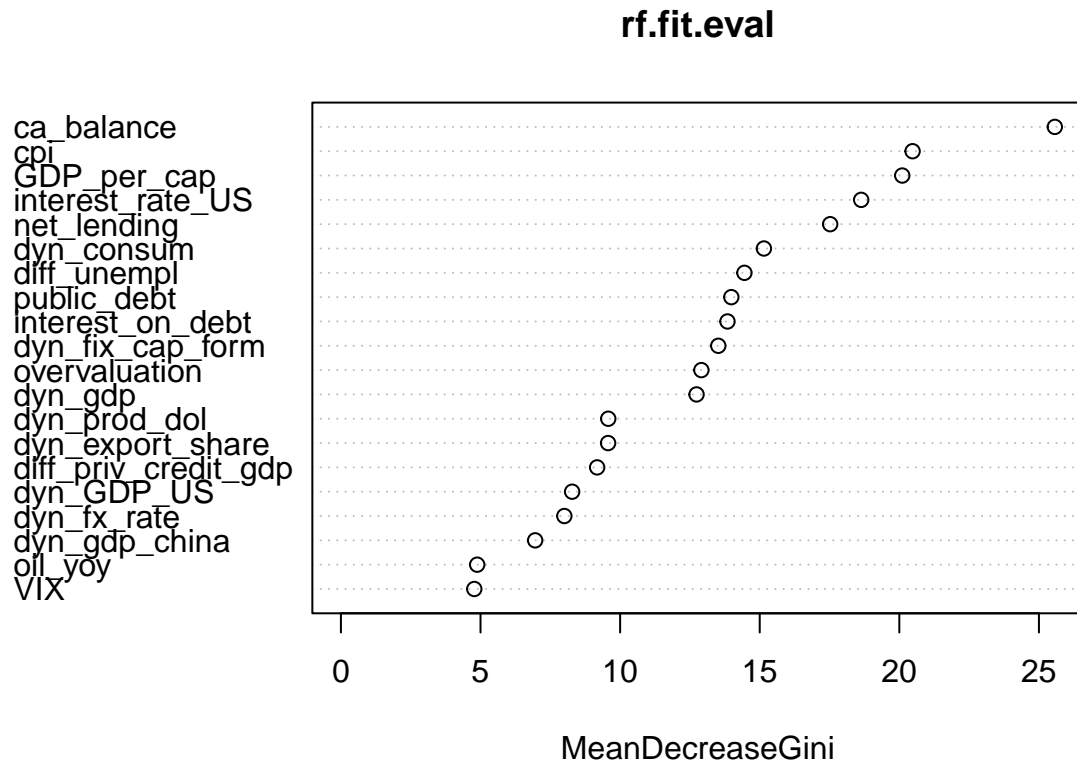
5.1 Performance

	Logit LASSO		Random Forest	
	advanced dummy	GDP per capita	advanced dummy	GDP per capita
% of correctly classified stress episodes	86.71	73.36	88.94	91.02
% of correctly classified tranquil episodes	50.67	72.47	66.62	66.04
Average	68.69	72.92	77.78	78.53
AUROC	0.83	0.85	0.88	0.89

Table 1: Average prediction accuracy of early warning models for years 2009-2018 (all observations used)

5.2 Interpretability

5.2.1 Variable Importance



5.2.2 Shapley Values

5.2.3 Partial dependence plots

5.2.4 Accumulated local effects plots

6 Conclusion

7 References

- Hastie, Trevor, Robert Tibshirani, Jerome H Friedman, and Jerome H Friedman. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Vol. 2. Springer.
- Jarmulska, Barbara. 2020. “Random Forest Versus Logit Models: Which Offers Better Early Warning of Fiscal Stress?” *ECB Working Paper Series No 2408 / May 2020*.