

Essays on Applied Machine Learning

Chapter 4

Falco J. Bargagli Stoffi

Machine Learning for Zombie Hunting

In this contribution, we exploit machine learning techniques to predict the risk of failure of firms. Then, we propose an empirical definition of *zombies* as firms that persist in a predicted status of high risk – i.e., beyond the highest failure risk decile – after which we observe that the chances to transit to lower risk are minimal.

We implement a Bayesian Additive Regression Tree with Missing Incorporated in Attributes (BART-MIA), which is specifically useful in our setting as patterns of undisclosed accounts correlate with firms' failures. We argue that these missingness patterns are likely to be generated from less viable firms being more prone to undisclosed their accounts. Thus, the inclusion of information on missingness patterns in the accounts is deemed important to boost the signal inputted in the model for predicting firms' viability.

After training our algorithm on 304,906 firms active in Italy in the period 2008-2017, we show how the proposed machine learning method outperforms proxy models like the Z-scores and the Distance-to-Default, traditional econometric methods, and other widely used machine learning techniques. These results provide evidence that economic intuition on the underlying data generating process can inform the choice of the predictive algorithm and can lead to significant improvements in the model's performance.

Lastly, we document that *zombies* are on average 21% less productive, 76% smaller, and they increased in number in times of financial crisis. In general, we argue that our application helps in the design of evidence-based policies in the presence of market failures, for example optimal bankruptcy laws. We believe our framework can help to inform the design of support programs for highly distressed firms after the recent pandemic crisis.