MACHINE LEARNING METHODS TO EXAMINE FINANCIAL MISREPORTING

[1]ARMAN HASSANNIAKALAGER, [2]PIETRO PEROTTI, [3]FANIS TSOLIGKAS

School of Management, University of Bath
E-mail: p.perotti@bath.ac.uk

Abstract - This paper introduces a new fraud detection model to the accounting literature using machine learning (ML). This model, which we refer to as Log it Boost, applies ensemble learning to logistic regressions. We show, using seven alternative measures assessing the ability to detect fraud, that our model outperforms the methods based solely on logistic regressions or other ML methods used by prior literature. Additionally, our model outperforms the others in predicting fraud beyond the current accounting period. Importantly, our method relies on a lower number of predictors than those used in prior ML research, thus minimizing concerns over multicollinearity and potential overfitting associated with machine learning methods.

JEL Classification: C44; C50;C53

Keywords - Machine Learning; Logistic Regressions; Accounting Irregularities; Aaers