

Asian Conference on Intelligent Information and Database Systems, 3-5 April 2017, Kanazawa, Japan

Increasing the detection of minority class instances in financial statement fraud

Stephen Obakeng Moepya^{1,2(B)}, Fulufhelo V. Nelwamondo^{1,2}, and Bhekisipho Twala²

¹ CSIR Modelling and Digital Science, Pretoria, South Africa
{Smoepya, fnelwamondo}@csir.co.za

² Faculty of Engineering and the Built Environment,
University of Johannesburg, Johannesburg, South Africa
btwala@uj.ac.za, <http://www.csir.co.za>

Abstract

Financial statement fraud has proven to be difficult to detect without the assistance of data analytical procedures. In the fraud detection domain, minority class instances cannot be readily found using standard machine learning algorithms. Moreover, incomplete instances or features tend to be removed from investigations, which could lead to greater class imbalance. In this study, a combination of imputation, feature selection and classification is shown to increase the identification of minority samples given severely imbalanced data.