

Bond University  
Research Repository



## Bankruptcy Prediction of Australian SMEs and Large Companies

Halteh, Khaled; Gepp, Adrian; Kumar, Kuldeep

*Licence:*  
CC BY-NC-ND

[Link to output in Bond University research repository.](#)

*Recommended citation(APA):*  
Halteh, K., Gepp, A., & Kumar, K. (2020). *Bankruptcy Prediction of Australian SMEs and Large Companies*. Abstract from International Online Conference in Applied Statistics 2020: Application of Statistics in Sciences, Social Sciences, Commerce, Humanities and Management , Mumbai, India.

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

For more information, or if you believe that this document breaches copyright, please contact the Bond University research repository coordinator.

## **ABSTRACT**

**Purpose:** This paper outlines the intrinsic differences between SMEs and large companies and develops bankruptcy prediction models using machine learning techniques to showcase the most important variables and the superior technique at predicting bankruptcies.

**Methodology:** Financial data pertaining to companies listed on the ASX were extracted and subsequently segregated using established metrics to distinguish SMEs from large companies. To overcome the class imbalance problem, Synthetic Minority Oversampling Technique (SMOTE) was used. After this, several machine learning techniques, including: decision trees, random forests, and stochastic gradient boosting, were applied to the datasets.

**Findings:** The models using SMOTE outperformed the original data across all machine learning models. Stochastic gradient boosting was found to be the superior model at classifying successful versus unsuccessful companies. The three most important variables for predicting bankruptcies of SMEs were: return on capital, cash divided by total assets, and return on equity. The large companies' three most important variables were: the natural logarithm of employee numbers, gross margin, and current assets divided by current liabilities.

**Research Implications:** The findings of the study confirmed the superiority of using SMOTE versus original data when there is a class imbalance problem. Both visual and empirical findings were used to showcase the most accurate classifying technique. Lastly, the study showcased that the most important variables pertaining to bankruptcy prediction differ amongst SMEs and large companies.

**Practical Implications:** This paper presents stakeholders, in general, and managers, in particular, with valuable information pertaining to the variable differences affecting financial health between SMEs and large companies. These can be used to aid decision-making processes to when determining the financial status of companies, thus can lead to the implementation of preemptive measures to prevent potential bankruptcies.

**Originality:** The SMOTE-machine learning approach used in this study is novel. No studies – to the best of the authors' knowledge – applied the aforementioned methodology to ASX companies.

**Keywords:** SMEs; Large Companies; Decision Trees; Random Forests; Stochastic Gradient Boosting; SMOTE; Class Imbalance