

A HYBRID APPROACH OF DEA, ROUGH SET THEORY AND RANDOM FORESTS FOR CREDIT RATING

DER-JANG CHI¹, CHING-CHIANG YEH^{2,*} AND MING-CHENG LAI³

¹Department of Accounting
Chinese Culture University

No. 55, Hwa-Kang Road, Yang-Ming-Shan, Taipei City 11114, Taiwan
dichi@ms43.hinet.net

²Department of Business Administration

³Graduate Institute of Business Administration
National Taipei College of Business

No. 321, Sec. 1, Jinan Road, Zhongzheng District, Taipei City 100, Taiwan

*Corresponding author: ychinc@webmail.ntcb.edu.tw; laimc@mail.ntcb.edu.tw

Received April 2010; revised September 2010

ABSTRACT. *In recent years, credit rating analysis has attracted lots of research interest in the literature. While the operating efficiency of a corporation is generally acknowledged to be a key contributor to the corporation's risk, it is usually excluded from early prediction models. To verify the operating efficiency as predictive variables, we propose a novel model to integrate rough set theory (RST) with the random forests (RF) technique, in order to increase credit rating prediction accuracy. In our proposed method, data envelopment analysis (DEA) is employed as a tool to evaluate the operating efficiency. Furthermore, the RST approach is used for variable selection due to its reliability in obtaining the significant independent variables, and utilized as a preprocessor to improve credit rating prediction capability by RF. The effectiveness of this methodology is verified by experiments comparing the RF, and compares the accuracy of the same prediction method with and without the DEA variable. The results show that operating efficiency does provide valuable information in credit rating predictions and the proposed approach provides better classification results.*

Keywords: Credit rating, Rough set theory, Random forests, Data envelopment analysis

1. Introduction. Credit ratings have been extensively used by bond investors, debt issuers, and governmental officials as an estimate of credit condition or the ability to pay debt. They are important determinants of risk premiums and even the marketability of bonds [1]. Numerous useful techniques for corporate credit rating prediction have been the subject of research in the academic and business community. However, whilst these are well-established credit rating prediction techniques, two main problems arise.

Firstly, early studies considered only financial factors as independent (input) variables [2-4]. Although financial factors, originally found in corporation's financial statements, can reflect some characteristics of a corporation from various aspects, the operating inefficiency of a corporation is also acknowledged to be a key contributor to a corporation's operation risk [5-8], and however, it is usually excluded from early prediction models. In this study, we believe operating efficiency, which reflects the status of the management of a corporation in credit ratings prediction, is a decisive factor affecting prediction accuracy. It is difficult to evaluate the efficiency of a corporation directly from its financial statements. An approach known as data envelopment analysis (DEA) may offer useful