

Corporate Credit Rating Estimation Using Stacking with Base Models Based on Groups of Classified Features

Sho Sato

Department of Industrial Engineering and Management,
Graduate School of Engineering
Kanagawa University
Kanagawa, Japan
r202370138yo@jindai.jp

Hideki Katagiri

Department of Industrial Engineering and Management,
Faculty of Engineering
Kanagawa University
Kanagawa, Japan
katagiri@kanagawa-u.ac.jp

Masanobu Matsumaru

Research Institute for Engineering
Kanagawa University
Kanagawa, Japan
pt120916cz@kanagawa-u.ac.jp

Abstract

A corporate credit rating is an index that evaluates a company's ability to fulfill its debt obligations on a 22-point scale from AAA to D. Since obtaining corporate credit rating is very costly, corporate credit rating estimation using machine learning has attracted much attention in recent years. The problem with conventional research is that certain features become noise in the estimation of specific companies, resulting in localized loss of accuracy. In the previous studies, the overall accuracy was improved by using new features in addition to financial indicators in the estimation of corporate credit ratings. However, because the new features became noise in the estimation, the accuracy of some companies was lower than that of the corporate rating estimation using only financial indicators as the features. On the other hand, the stacking suppresses the accuracy loss due to noise. The stacking is built in two stages: in the first stage, the base model is built, and in the second stage, the meta-model is built using the estimates of the base model as features. The design of the base models is critical to the accuracy of the model with the stacking. In this study, we design the base models based on the groups of classified features. For example, the groups of features related to financial indicators includes profitability and safety, etc. The objective of this study is to construct a proposed model using the stacking with base models based on the groups of classified features. In addition, to confirm the usefulness of the proposed model, we compare the accuracy of the proposed model with that of the conventional model that does not use stacking. In this study, the corporate credit rating estimation is treated as a 22-class multi-label classification problem. The method used in the corporate credit rating estimation model is the Light Gradient Boosting Machine. In this study, the proposed model uses rating estimates of seven base models. The seven base models include a conventional model, five models using financial indicators, and a model using investment and loan network indicators. The financial indicators used in the five models are based on five groups of financial indicators classified by Nikkei NEEDS-FinancialQUEST. The five groups are profitability, return on capital, cash flow, return on equity, and safety. We used corporate ratings from Japan Credit Rating Agency, Inc. for the period from 1999 to 2021 as the labeled training data. Numerical experiments were conducted using data from 1,972 firms with 167 indicators. As a result, the proposed model improved the estimation of 53 firms compared to the previous model. The accuracy of the proposed model is 0.798, which is 0.027 points higher than that of the conventional model. The Quadratic Weighted Kappa is

0.914, which is 0.012 points higher than that of the previous model. The accuracy of the proposed model was better than that of the conventional model, indicating the usefulness of the proposed model.

Keywords

Corporate credit rating, Machine learning, Stacking and Multi-label classification.

Biographies

Sho Sato is with Department of Management Engineering, Graduate School of Engineering, Kanagawa University. He is researching on corporate credit risk analysis using machine learning. He entered Kanagawa University in 2019 and joined the Management Systems Engineering Laboratory in 2021. After graduating from the Department of Management Engineering in 2023, he entered the graduate school of Kanagawa University. At the Mathematical Systems Academic Conference 2021, held online on February 10, 2021. He made a presentation and received the "NTT Data Mathematical Systems Student Research Encouragement Award for Excellence 2021" at the Mathematical Systems Academic Conference held online on February 10, 2021.

Hideki Katagiri is a Professor of Department of Industrial Engineering and Management at Kanagawa University, Japan. He earned his B.E., M.E. and Ph.D. in Engineering at Osaka University in 1995, 1997 and 2000, respectively. He was the Chair of IEEE SMC Hiroshima Section Chapter (2008-2010) and a Visiting Scholar at the University of Chicago Booth School of Business (2014-2015). He was a Visiting Professor of Hiroshima University (2016-2020). His research and teaching activities are in the areas of operations research and soft computing, especially, multi-objective optimization under uncertainty and data analysis using machine learning techniques. He is the author or co-author of more than 100 refereed journal papers and several co-authored or co-edited books in English.

Masanobu Matsumaru is currently a visiting professor in Research Institute for Engineering at Kanagawa University, and was a former professor of Department of Industrial Engineering and Management at Kanagawa University, Japan. He earned Bachelor of Engineering in Faculty of Science and Engineering from Waseda University, Japan, Master of Engineering in Master course Graduate school of Science and Engineering from Waseda University and Doctor of Engineering from Tokai University. His research interests include Industrial Engineering and Management, Management Modeling, Problem Solving, Financial Management and Supply Chain.