Corporate Rating Model using Threshold Optimization and XGBoost

Anda Zhang
Department of Industrial Engineering and Management
Graduate School of Engineering, Kanagawa University
Kanagawa, Japan
r201970101qa@jindai.jp

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

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

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

Abstract

This paper proposes a corporate rating model by eXtreme Gradient Boosting (XGBoost), which is one of the machine learning algorithms. The rating is a graded evaluation as the safety of the ability to pay principal and interest and is the opinion of the rating agency. In a situation where the rating components of the rating agency have not been published, it is considered meaningful study to perform the rating if the estimated rating scores are close to the rating score of the rating agency using financial data. Ratings consist of a small number of different grade classes, labeled by using alphabet character strings like "AAA". This model incorporates Quadratic Weighted Kappa (QWK), which is an evaluation index that considers the order relationship of rating classes and the optimization of class thresholds. By optimizing the class delimiter, which is the threshold of the QWK evaluation class, the accuracy of the rating model is improved. As the result of comparing with the rating score of the rating agency called Rating and Investment Information, Inc. to verify the validity of the rating of the model, it was confirmed that the rating score was able to approach the rating of the rating agency.

Keywords

Corporate rating, eXtreme Gradient Boosting (XGBoost), Rating agency, Financial data, Class threshold optimization, Quadratic Weighted Kappa(QWK), Empirical analysis

1. Introduction

Ratings are an evaluation of the safety of the ability to pay principal and interest, and are the opinion of rating agencies. Ratings are also used by stakeholders around companies as one of the means to evaluate corporate creditworthiness. These stakeholders include business partners, lenders, investors such as shareholders, employees, consumers, countries and public bodies. The risks surrounding companies in today's complex and rapidly changing times are enormous, and companies operating under such circumstances want to reduce the risks as much as possible. Ratings are used not only when companies issue corporate bonds, but also as a method for estimating the risk of trading companies even for companies that do not issue them, and are also one of creditworthiness evaluations. In addition, disclosure materials such as securities registration statements and prospectuses submitted by companies when issuing corporate bonds are required to include a rating. It is stipulated by law that these ratings must belong to "designated rating agencies". This designated rating agency is stipulated by the Cabinet Office Ordinance on Disclosure of Corporate Information, etc. (January 30, 1973, Ministry of Finance Ordinance No. 5), Article 1, Item 13-2, and is published by the Disclosure Office Ordinance. In Japan, there are only two institutions, rating and Investment Information, Inc. and Japan Credit Rating Agency, Inc. Therefore, since the ratings of these two rating agencies have a certain degree of evaluation, when a company wants to evaluate the creditworthiness of a trading company, it is possible to request a rating from these two rating agencies to obtain a rating.

Companies want to ask a rating agency to get a rating, however, the cost of the request is so high that companies with a weak financial base cannot use the rating agency's rating. If there is a method that can obtain a rating as close as possible to the rating of the rating agency from the published financial indicators, it is expected that the company will be rated by the method proposed in this study. The cost is so high that companies with a weak financial base cannot use the valuation for rating agencies. If there is a method that can obtain a rating as close as possible to the rating of the rating agency from the published financial indicators, it is expected that this method will be used to rate the trading company. Against this background, the study proposes a method for estimating a rating close to that of a rating agency from the financial data published in this study. The rating agency's rating component is not publicly available. Under these circumstances, it would be meaningful to use financial data to perform an evaluation when it is close to the rating agency's rating. Ratings consist of a small number of different grade classes, usually labeled by using alphabet character strings like "AAA". This paper proposes a corporate evaluation model by eXtreme gradient boosting (XGBoost), which is one of the machine learning algorithms, using financial data. When performing an actual rating, the major issue is which rating class to rate. In this study, Quadratic Weighted Kappa (QWK) is used as an evaluation index that favors the order relation between these multiple classes. QWK is improved by optimizing the rating class threshold (class threshold). It is compared with the rating agency's rating to validate the evaluation of the proposed model. In addition, many financial indicators are selected by the Gini index, one of XGBoost's variable selection methods, as the optimal variables in the rating, and are manifested as the importance of the variables that explain the rating.

2. Literature Review

XGBoost is one of the gradient boosting using decision trees and used by the top teams in many Kaggle machine learning competitions (2016). XGBoost has been successfully applied in real-life data of companies. Xia et al. proposed a sequential ensemble credit scoreing model based on XGBoost (2017). Zieba et al. has been successfully applied in bankruptcy prediction on real-life data of Polish companies (2016). We propose a rating model using XGBoost. As a result of examining conventional studies, there are rating studies in Japan. However, a rating model using XGBoost has not been proposed. Yamamoto (2016) constructed a rating model using an ordinal logit model. Tanaka et al. (2014) proposed a method for multigroup discriminant analysis based on Support vector machines (SVM) and performed comparative verification with the ordinal logit model and it said that SVM method for determining corporate rating is better than an ordinal logit model.

3. Methods

3.1 XGBoost

XGBoost is a type of gradient boosting using decision trees. Boosting is ensemble learning derived from constructing a strong learner by combining weak learners. Gradient boosting uses the gradient to minimize the prediction results

up to the previous weak learner and the loss function of the true objective variable. XGBoost weights the leaf nodes using multiple parallelized decision trees and calculates predicted values using a pseudo-objective function.
Now let D = {$\boldsymbol{x}$; y} be the training dataset set for the input dataset $\boldsymbol{x}$ and the output dataset y. $\hat{\boldsymbol{y}}_{\boldsymbol{\iota}}$ is the ith predicted value at the bth boost. $\boldsymbol{f_b}$ is a score calculated using a tree structure with leaves j with a weight score w. The final predicted value $\hat{\boldsymbol{y}}_{\boldsymbol{\iota}}$ is calculated by summing the scores over all the leaves and is calculated by Eq. (1).

$$\hat{y}_l = \sum_{b=1}^{B} f_b\,(x_i) \qquad (1)$$

Next, the gradient is used to minimize the loss function Lb defined in Eq. (2).

$$L_b = \sum_{i=1}^{n} l(y_i, \hat{y}_l) \qquad (2)$$

Eq. (3) adds a regularization term to the defined loss function to minimize the loss function.

$$L_b^X = L_b + \gamma T + \frac{1}{2}\lambda \|w\|^2 \qquad (3)$$

Here, γ is a hyper parameter of minimum loss reduction for further division, and λ is a parameter to prevent overfitting. $\|\boldsymbol{w}\|^{\mathbf{2}}$is the weight of the leaves in the $\boldsymbol{L_b}$ norm, and T is the number of leaves in the tree.

3.2 Quadratic Weighted Kappa as an evaluation index

The ordering of rating classes is taken into account when building the model. Quadratic Weighted Kappa is used as a rating index. This index is designed so that the larger the prediction, the greater the penalty. This guarantees the consistency between the prediction class estimated by the model and the true rating class of the rating agency, and also guarantees the closeness of the rating classes of both. In other words, it represents the degree of agreement between the classification result of the predicted value by the model and the classification result of the true value, and is expressed by Eq. (4).

$$\kappa = 1 - \frac{\sum_{i,j} \alpha_{i,j} Q_{i,j}}{\sum_{i,j} \alpha_{i,j} E_{i,j}} \qquad (4)$$

Here, $Q_{i,j}$ is the number of records whose true value class is i and whose predicted value class is j, and when written in the form of a matrix, it becomes a multiclass confusion matrix. $E_{i,j}$ is the expected value of the number of records belonging to each cell (i, j) of the confusion matrix i, j when the distributions of the true value class and the predicted value class are independent of each other. It is calculated as "(the ratio of the true value) × (the ratio of the predicted value j) × (the number of records in the entire data)". $\alpha_{i,j}$ is expressed by the square of the difference between the true value and the predicted value $(i - j)$^2. If the prediction value j is far away from the true value i of the rating, the penalty will be large. This QWK takes a value from -1 to 1; κ = 1 when the true value i and the predicted value j completely match, and κ = -1 when the true value i and the predicted value j do not completely match, κ = 0 when they unrelated.

3.3 Optimization of rating class thresholds in Quadratic Weighted Kappa

QWK is a useful index for multi-classes that have an ordinal relationship, but it is not possible to obtain a good score simply by outputting the predicted class as it is. Therefore, an approach is used in which the threshold value between classes is optimized and calculated after the predicted value is output as a continuous value. Specifically, when the rating class s and the probability of being rated as belonging to the rating classes are $\boldsymbol{p_s}$, $\sum_{\boldsymbol{s}} \boldsymbol{p_s}$ is the predicted value. The procedure after this is to calculate the optimization to find the delimiter value for classification.

QWK can correct large deviations, but simply setting the threshold between classes to 0.5 will still result in a large number of small deviations. Therefore, it is necessary to adjust. Table 1 shows the ratings before adjusting the interclass threshold.

Table 1. Before adjusting the interclass threshold

	AA-	A+	A	A-	BBB+	BBB
AA-	12	2	1	1	0	0
A+	3	25	10	0	0	0
A	1	2	43	6	0	0
A-	0	0	10	34	1	0
BBB+	0	0	0	8	9	1
BBB	0	0	1	6	0	8

On the other hand, when the threshold was adjusted and optimized, the number of small deviations decreased as shown in Table 2. From this, it was confirmed that optimizing the threshold between classes is important.

Table 2. After adjusting the interclass threshold

	AA-	A+	A	A-	BBB+	BBB
AA-	24	2	0	0	0	0
A+	3	28	3	0	0	0
A	0	1	45	4	0	0
A-	0	1	4	31	0	0
BBB+	0	0	1	2	14	3
BBB	0	0	1	4	2	6

4. Data Collection

The rating score of Rating Agency R&I are used as a supervised benchmark to verify the validity of the rating of the proposed model. 149 types of financial indicators of companies listed on the Tokyo Stock Exchange are used as the explanatory variables of rating. The analysis period is 21 years from 2000 to 2020, and the cumulative number of data is 3636 companies including manufacturing industry 2,145 and non-manufacturing industry 1,491 companies. Financial data is obtained from Nikkei NEED-Financial QUEST.

5. Results and Discussion

5.1 Numerical Results

Table 3 shows the results of the accuracy and QWK of this study. As a result, the QWK introduced to optimize the rating class threshold (class delimiter), which was the purpose of this study, was close to 1, achieving the purpose and obtaining accurate results did it.

Table 3. Accuracy and QWK

Industry	Accuracy	QWK
Manufacturing industry	0．855	0．957
Non-manufacturing industry	0．821	0．937
All industries	0．872	0．968

5.2 Comparison with using QWK and non-using QWK

From the simulation results shown in table 4 in the rating model that does not use QWK, there are cases where the 'AAA 'rank is judged to be 'B' rank and the 'B' rank is judged to be 'AAA' rank. There is a large discrepancy between the prediction class and the actual class.

Table 4. Demonstration analysis results without QWK

		Predicted value class						
		AAA	AA	A	BBB	BB	B	CCC
True value class	AAA	33	7	2	0	0	20	0
	AA	11	20	5	0	0	2	0
	A	0	11	47	0	0	0	0
	BBB	0	0	0	13	15	0	2
	BB	0	0	0	5	68	13	0
	B	16	2	2	0	20	67	0
	CCC	0	0	0	3	0	0	10

On the other hand, as shown in Table 5, optimizing the hyper parameters using QWK corrects this large deviation, and the predicted rank value of the rating approaches the true rank value. As a result, a high value of QWK = 0.968 was obtained. Therefore, the significance of using QWK, which is a feature of this study, was confirmed.

Table 5. Demonstration analysis results when using QWK

		Predicted value class											
		AAA	AA+	AA	AA-	A+	A	A-	BBB+	BBB	BBB-	BB+	BB-
True value class	AAA	5	0	0	0	0	0	0	0	0	0	0	0
	AA+	0	24	0	0	0	0	0	0	0	0	0	0
	AA	0	0	25	1	0	0	0	0	0	0	0	0
	AA-	0	0	0	39	1	0	0	0	0	0	0	0
	A+	0	0	0	3	31	7	3	1	0	0	0	0
	A	0	0	0	0	3	62	10	0	0	0	0	0
	A-	0	0	0	1	0	5	68	2	0	0	0	0
	BBB+	0	0	0	0	1	1	4	29	1	0	0	0
	BBB	0	0	0	0	1	0	1	3	21	0	0	0
	BBB-	0	0	0	0	0	0	0	0	1	8	0	0
	BB+	0	0	0	0	0	0	0	0	0	0	2	0
	BB-	0	0	0	0	0	0	0	0	0	0	0	1

5.3 Comparison of important explanatory variables

Figures show the important explanatory variables clarified by the Gini coefficient of the variable selection method of XGboost for manufacturing, non-manufacturing, and all industries. It was found that the important explanatory variables in manufacturing are different from non-manufacturing ones. In manufacturing industry, the important explanatory variables are capital indicators like the net working capital amount, net working capital and investment indicators like cash flow to capital investment ratio and depreciation rate.

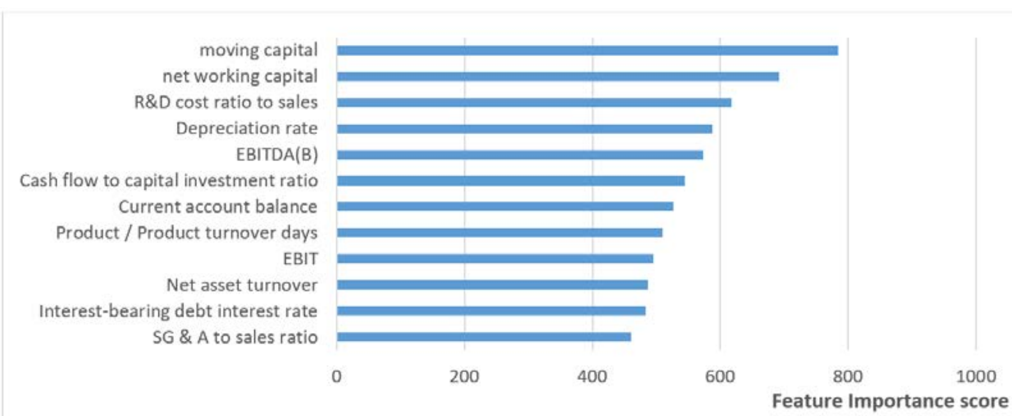

Figure 1. Feature Importance (manufacturing industry)

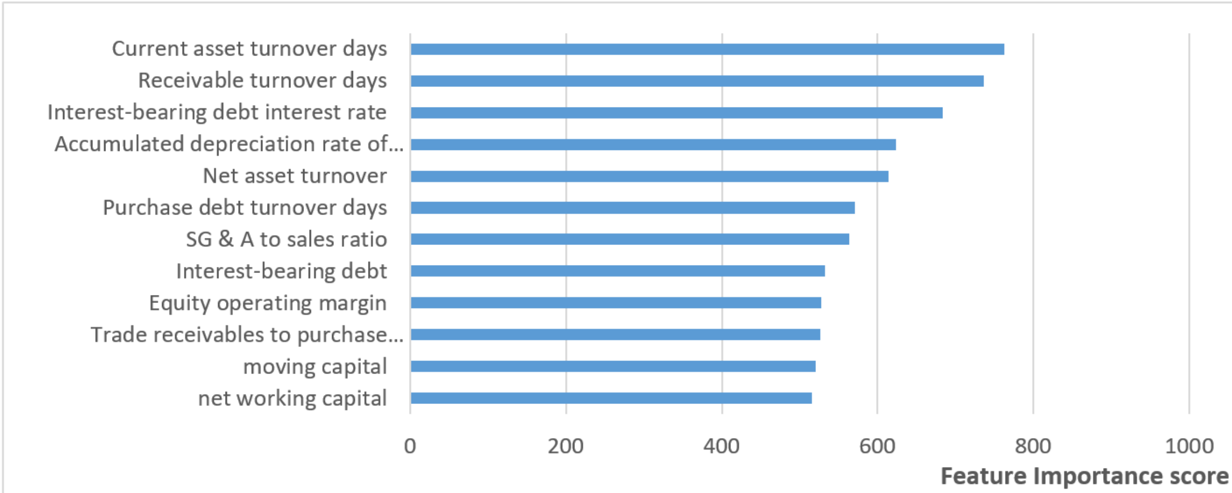

Figure 2. Feature Importance (non-manufacturing industry)

In the non-manufacturing industry, the important explanatory variables are profitability indicators like current asset turnover days and trade receivable turnover days and the safety indicators like interest-bearing debt interest rate, property, plant and equipment depreciation and net asset turnover. These results indicate that the important explanatory variables for rating are clearly different between the manufacturing industry and the non-manufacturing industry.

In all industries, the important explanatory variables are cash flow indicators like EBITDA, profitability indicators like SGA ratio and safety indicators like the net working capital amount in a well-balanced manner.

Next we will explain the reasons for companies whose ratings are significantly different from those of rating agencies. As the first company, we focused on Taiyo Nippon Sanso, which is a company with a large divergence, with the rating agency rating being BBB, while the estimation is A + rating. Examining the cause, the net working capital amount has improved significantly from 60,975 million yen to 292,480 million yen. The net working capital amount is the funds invested amount in business activities and is the safety indicators. This amount improves the rating significantly.

Mitsui Chemicals has been downgraded from A + to BBB +. The reason is that the number of days for raw materials and other rotations has almost doubled from 7.45 days to 14.76 days, and it can be seen that the rating evaluation was lowered due to the deterioration of the number of rotation days. Japan Airlines has improved ratings from A- to AA-. The reason is that the company's net working capital has improved from -201,572 million yen to -147,971 million yen.

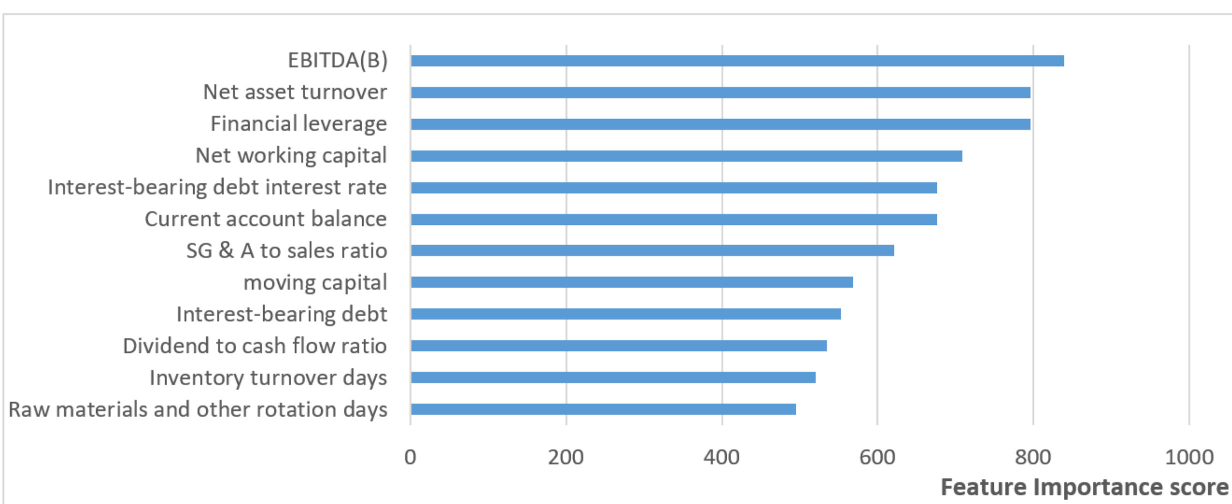

Figure 3. Feature Importance (all industry)

TEPCO's rating has risen from BBB + to A +. It can be seen that the reason is that the amount of interest-bearing debt has been evaluated to have decreased significantly from 6,022,968 million yen to 4,914,929 million yen. At the same time, the current account balance improved from 586,686 million yen to 1,004,878 million yen, and it became clear that the improvement in the profit structure was also reflected in the rating

5.4 Comparison of XGBoost and SVM

As mentioned in the previous research, Tanaka et al. [3] proposed the SVM method. It is reported that SVM showed results that generally exceeded the hit rate by rating and overall average than the ordinal logit model. Therefore, in this study, we compared it with SVM. The results are shown in Table 6.

Table 6 Comparison of XGBoost and SVM

	Accuracy	QWK
XGBoost	0.872	0.968
SVM	0.738	0.886

As can be seen from the above, XGBoost showed better values than SVM in both Accuracy and QWK. It shows its usefulness.

6. Conclusion

We proposed a corporate rating model using XGBoost, in which the evaluation index QWK is incorporated to take into account the ordering relationship of rating classes and class threshold optimization. As mentioned in the beginning, companies want to ask a rating agency to get a rating. However, the cost of the request is so high that companies with a weak financial base cannot use the rating agency's rating. Then the study proposed the model for estimating a rating close to that of a rating agency from the financial data published. The model evaluated the companies rating close to the rating agency's rating. Using this model, the stakeholders like business partners, lenders, investors such as shareholders, employees, consumers, countries and public bodies can use this model to rating the risk.

We also clarified the usefulness of using the evaluation index QWK. QWK was improved by optimizing the rating class threshold (class threshold) and was compared with the rating agency's rating to validate the evaluation of the proposed model. The important explanatory variables indicated that the important explanatory variables for rating were clearly different between the manufacturing industry and the non-manufacturing industry. For future research, we will plan to give a rating in consideration of factors other than financial indicators.

References

J. H. Friedman, Greedy function approximation a gradient boosting machine, *The Annals of Statistics*, Vol.29, No.5, pp.1189-1232, 2001

J. H. Friedman, Stchastic gradient boosting, *Computational Statistic & Data Analysis*, Vol.38, No.4: pp.367-378, 2002

J. H. Friedman, T.Hastie, and R. Tibshirani, Additive logistic regression: a statistical view of boosting, *Annals of Statistics*, Vol.28, No.2, pp. 337-407, 2000

J. H. Friedman and B.E. Popescu, Importance sampled learning ensembles, *Technical report, Dept. Statistics*, Stanford Univ., 2003

K. Tanaka, H. Nakagawa, Proposal of SVM method for determining corporate rating and verification of effectiveness by comparison with sequential logit model, *Transactions of the Operations Research Society of Japan*, Vol. 57, pp. 92-111 (2014) [in Japanese]

M. Matsumaru, T. Kawanaka, S. Kaneko, and H. Katagiri, Bankruptcy prediction for Japanese corporations using support vector machine, artificial neural network, and multivariate discriminant analysis, *International Journal of Industrial Engineering and Operations Management (IJIEOM)*, Vol.1, No.1, pp.78-96, 2019

M. Zieba, S. K. Tomczak, and J. M. Tomczak, Ensemble boosted trees with synthetic features generation in application to bankruptcy prediction, *Expert Systems with Applications*, Vol. 58, pp. 93-101, 2016.

R. Yamamoto, Credit Risk Analysis Beginning with R, *Operations Research*, Vol. 61, No. 6, pp. 365-370 (2016) [in Japanese]

T. Chen and C. Guestrin, XGBoost: A Scalable Tree boosting System, KDD '16: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.785–794, https://doi.org/10.1145
/2939672.2939785Cornel University, 2016

T. Chen and T. He., Xgboost: extreme gradient boosting, *Package Version: 0.90.0.2*, 2019

Y. Xia, C. Liu, Y. Li, and N. Liu, A boosted decision tree approach using Bayesian hyper parameter optimization for credit scoring, *Expert Systems with Applications*, Vol. 78, pp. 225- 241, 2017.

Biography

Anda Zhang is graduate student of Industrial Engineering and Management Department, Kanagawa University, Japan. His research interest contains algorithms and applications of machine learning techniques, especially eXtreme Gradient Boosting (XGBoost), in financial areas. He has been studied bankrupt modelling using machine learning techniques already. His research interests algorithms and applications of XGBoost in financial areas. His current research is to construct the corporate rating models to help investors reduce costs.

Hirohisa Hirai is a Professor of Department of Industrial Engineering and Management at Kanagawa University, Japan. He earned his Ph.D. in Engineering at Osaka University respectively. His research is an empirical study in the areas of management accounting, firm analysis and applied statistics, especially, firm valuation. He has received paper awards at several academic societies in the areas of accounting and management.

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.

Hideki Katagiri is a Professor of Department of Industrial Engineering and Management at Kanagawa University, Japan. He is a Visiting Professor of Hiroshima 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). 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.