

## An Approach for Academic Success Predictive Modeling based on a Multi-objective Genetic Algorithm

Nobuhiko Kondo <sup>\*</sup>, Takeshi Matsuda <sup>\*</sup>, Yuji Hayashi <sup>\*</sup>,  
Hideya Matsukawa <sup>†</sup>, Mio Tsubakimoto <sup>‡</sup>, Yuki Watanabe <sup>§</sup>,  
Shinji Takeishi <sup>\*\*</sup>, Hideaki Yamashita <sup>\*</sup>

### Abstract

In this paper, we propose a new method of constructing machine learning models for predicting academic success. In this method, a multi-objective genetic algorithm is deployed to select explanatory variables for the predictive model as an approach that takes into account both elements of predictive performance and interpretability. By using two evaluation functions, i.e., prediction performance and the number of explanatory variables, our method can find the Pareto-optimal solution set that reflects these trade-offs. The numerical simulation results show that our method can obtain a model set that takes into account the trade-off between the accuracy and complexity of the predictive model, although there are differences in behavior depending on the academic success indices to be predicted.

*Keywords:* academic success, predictive model, multi-objective genetic algorithm, variable selection, interpretability.

### 1 Introduction

Institutional research (IR) is a function that supports decision-making at universities based on salient data, and its importance has been increasingly recognized in recent years. In particular, IR is attracting increasing attention because of the growing availability of data due to the development of ICT, as well as the practical application of advanced statistical methods such as machine learning and data mining. In terms of evidence-based educational improvement, there has also been rapid development within paradigms such as learning analytics and educational data mining [1], and IR needs to collaborate with these fields to provide more advanced decision-making and learning support.

---

<sup>\*</sup> Tokyo Metropolitan University, Tokyo, Japan

<sup>†</sup> Tohoku University, Miyagi, Japan

<sup>‡</sup> The University of Tokyo, Tokyo, Japan

<sup>§</sup> Tokyo University of Science, Tokyo, Japan

<sup>\*\*</sup> University of Tsukuba, Ibaraki, Japan

In recent years, the assessment and visualization of learning outcomes have also been required from the perspective of quality assurance and the social responsibility of universities. The universities are now faced with the task of assessing learning outcomes appropriately according to a given university's diploma policy through IR and improving subsequent education based on this assessment. In this context, there is a great deal of research regarding a concept called student success, which describes the student's success status in some sense [2–4]. Student success does not necessarily refer only to the academic outcomes per se provided by a university or degree program but encompasses success in a broader sense, including community building, satisfaction, and career development. In this study, we refer to a student's achievement of some kind of academic success throughout their college life as “academic success” (AS), and we discuss the prediction of AS based on IR data for academic advising. Since the graduation rate of higher education institutions in Japan is over 90% [5], it is important to capture and evaluate AS including qualitative academic outcomes, not just completion and satisfaction. In the experiments described in this paper, we used data from students at a medium-sized university, to examine the results in a typical Japanese higher education institution.

The recent increasing importance of AS also requires universities to provide academic advice to guide students to achieve this AS. In particular, supporting data-driven AS advice is also a challenge for IR. Related to this issue, in the fields such as learning analytics mentioned above, the use of predictive models such as machine learning in learning and teaching has been actively studied [6]. The authors have also developed a prototype version of an IR system called “Risk Detector” for early detection of the possibility of a student dropping out and its use in the advising process [7]. According to the formative evaluation of the system, although the prediction accuracy was acceptable, it was indicated that not only the prediction accuracy but also the information on which variables are important for preventing students from dropping out of school is necessary for guidance. In addition, from a practical standpoint, a general-purpose system that can handle AS in a broader sense than just the perspective of retention per se is needed. Thus, to use predictive models to support the advising process for AS, it is simultaneously necessary to have not only high predictive performance but also the deployment of simple models with few explanatory variables. However, since there is sometimes a trade-off between model complexity and predictive performance [8], a model optimization method that simultaneously takes both requirements into account is needed.

Therefore, in this study, as the fundamental technology of the system, we propose an approach for the construction of an AS predictive model that automatically selects important explanatory variables useful for the advising process. The variables are selected using a multi-objective genetic algorithm. The predictive performance of AS and the model interpretability are considered simultaneously regarding the practicality for student advising. The proposed method is also designed to be generic so that it is applicable to a variety of AS indices. To verify that the proposed method is capable of constructing a model that takes into account the trade-off between prediction performance and model complexity, we performed numerical simulations using actual student data to investigate the behavior of the proposed method in practice.

This paper is organized as follows. In Chapter 2, we describe machine learning, genetic algorithms and their applications to education. In Chapter 3, we propose a method for constructing an AS predictive model using a multi-objective genetic algorithm. In Chapter 4, we seek to verify the method via numerical experiments, and Chapter 5 concludes the paper.

## 2 Machine Learning, Genetic Algorithms, and their Educational Applications

We apply a combined approach of machine learning and genetic algorithms in this study to build a predictive model of AS. In this chapter, we will outline machine learning and the genetic algorithm and briefly summarize the current status of their use in the field of educational data analysis.

### 2.1 Machine Learning and Predictive Models

Machine learning is a framework for building mathematical models of objects based on data. Recently, it has been actively studied as a fundamental technology for artificial intelligence, and its applications within various fields have been promoted. In particular, supervised learning approaches have been used for prediction and classification. Due to the breadth of its application field, it is frequently used. Based on multivariate data with teacher labels, supervised learning is used to estimate the model parameters statistically to make prediction and classification errors as small as possible. Hence, it is a statistical method for estimating the mapping from multidimensional explanatory variables to objective variables.

In learning analytics, an approach based on some kind of prediction using a predictive model is often used [6]. Machine learning methods can be used to build predictive models, as the size of data sets and the complexity of the subject matter increase. Many studies in the field of IR have been conducted on some kind of prediction based on IR data. There are numerous studies on predictions of retention and withdrawal [9, 10].

The study by Brooks and Thompson [6] summarizes the use of prediction models for learning analytics, namely, linear regression, logistic regression, nearest neighbor method, decision tree, naive Bayes classifier, Bayesian network, support vector machine, neural networks, and ensemble methods. In recent years, machine learning packages have become readily available in programming languages such as Python, and it is possible to build predictive models easily by machine learning through their use. In this study, we are considering predicting AS at the time of graduation at some point in the bachelor's program (e.g., at the end of the first year), which will be described later. Appropriate variables from the IR data available up to that point are used as explanatory variables when performing prediction by machine learning, and some indices indicating the appearance of AS are used as objective variables.

Logistic regression is used in this paper, although any machine learning algorithm can be used. Logistic regression is a kind of generalized linear model and is often used for two-class classification [11]. Because it often shows relatively good performance in addition to easy handling and interpretability, it is a model that is often used in real-world practice.

### 2.2 Genetic Algorithm

Evolutionary computation is a general term for computational methods that apply the genetic and evolutionary mechanisms characteristic of living organisms to learning and optimization processes. Evolutionary computation is a probabilistic multipoint search method that can identify multiple solutions in parallel, does not require gradient information of the objective function, has a mechanism to avoid local solutions by stochastic operations, and so on. It is known as metaheuristics that can find a practical solution in a realistic time for a complex real problem [12].

We use a genetic algorithm [13] in this study that is often used as one of the evolutionary computations. A genetic algorithm is a framework in which selection, crossover, and mutation are the main genetic operations. A solution candidate for an optimization problem is considered an individual in a genetic algorithm, and the population of multiple individuals is evolutionarily altered to look for the optimal solution as a population. Each individual holds the information of solution candidates as a symbol string such as a binary string, using the analogy of chromosomes. Each individual's fitness is based on the evaluation of the value of the objective function in the optimization problem. The fitness corresponds to how well the individual adapts to the environment. The population is altered by performing genetic operations such as selection, crossover, and mutation. Selection is an operation that determines the individuals to be found in the next generation, and is performed by such a method that the higher the fitness is found to be, the easier it is for the next generation to survive. After that, the chromosomes are exchanged (i.e., crossover), or the genes of the chromosomes are replaced (i.e., mutation) with a certain probability among the selected individuals, and the information is stochastically altered. By repeating these genetic operations, individuals that are suitable for the environment are generated as a group while partially inheriting the parental traits is simulated. Figure 1 shows the flow of this genetic algorithm. Genetic algorithms have good compatibility with NP-hard combinatorial optimization problems because they operate on symbol strings as solution candidates.

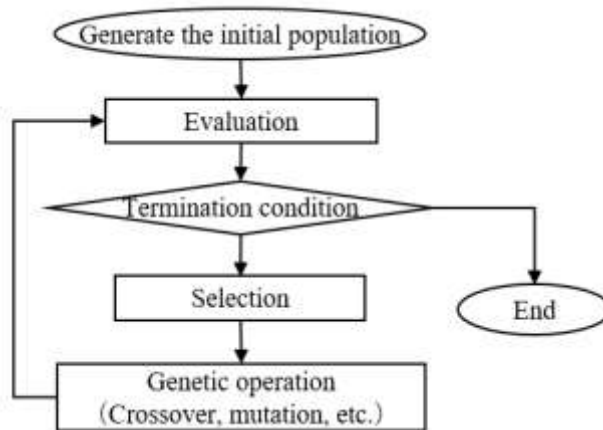


Figure 1: Evolutionary computation flowchart.

### 3 Proposed Method

We consider the early prediction of the degree of AS at the time of graduation in this study to use it for the guidance of students' AS. In this section, we propose a method for constructing an AS prediction model using machine learning and multi-objective genetic algorithms.

#### 3.1 Prediction of Academic Success by Machine Learning

We propose a general-purpose AS predictive model construction framework that can handle AS in the broader sense required in this paper. This framework considers predicting AS at graduation at some point in the ongoing bachelor's program (e.g., at the end of the first year). In this framework, the degree of AS is expressed quantitatively, and an AS index is set to predict the level of

this value. Because this is a typical classification problem, a machine learning (supervised learning) algorithm is used to construct the predictive model, as described in Section 2. From the IR data available up to the time of prediction, some salient variables are used as explanatory variables, and the AS indices are modeled as objective variables. Figure 2 illustrates this AS prediction model.

AS indices can be assumed at various levels such as universities and degree programs, for example, at the university level, the cumulative grade point average (GPA) at the time of graduation, the evaluation of generic skills defined in the diploma policy, or whether straightforward graduation can be considered. At the degree program level, evaluation of the specialized knowledge and skills set in the diploma policy for each degree program may be considered. Some of these AS indices can be obtained by using IR data held by the university, such as academic records and enrollment status, whereas other data can be obtained by indirect assessment by student survey or direct evaluation by performance assessment. The indices considered possible to be quantified by some method can be the target of prediction by this method.

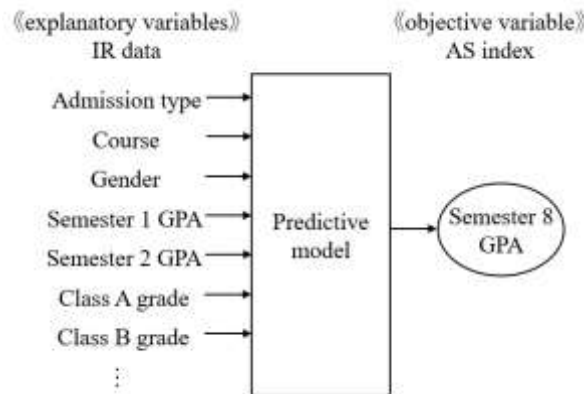


Figure 2: Predictive model for AS.

### 3.2 Selection of Explanatory Variables by Genetic Algorithm

It is considered possible to predict the degree of achievement of the AS index from the IR data by utilizing the predictive models using machine learning. However, it is difficult to give a guide of how to teach if the situation is such that only the prediction result of the index value is given, assuming that this is used for actual student guidance. It is thought that the degree of dependence on abilities and experience will increase.

The IR data, which is the explanatory variable of the predictive model, can be an important index for predicting AS or failure. However, the amount of IR data held by the university and available by linking to each student is increasing, in volume, and the explanatory variables that can be extracted as candidates from this are considered to be of higher dimensions.

It is difficult to establish which variables are relatively important when using a predictive model consisting of high-dimensional explanatory variables by simply showing the values of those explanatory variables. This is closely related to an operation called variable selection or feature

selection, which is generally important in statistical models. Because the prediction performance may deteriorate due to overtraining, it is important to select an appropriate subset of variables when the dimension of the explanatory variable of the model is high in regard to the training data. Moreover, the variable selection not only improves the prediction performance but also improves the interpretability of the model. Also, in the case of the proposed method, if the predictive model can be constructed as simply as possible by selecting only the important variables, the interpretability of the model can be expected to be increased in the sense that it shows the important variables regarding the degree of AS.

There are various methods for selecting such high-dimensional explanatory variables. We use a genetic algorithm in this study, which is one of the evolutionary calculations shown in Section 2. In the proposed method, we consider a chromosome consisting of as many binary values as candidate explanatory variables. Each binary value of the chromosome is associated with an explanatory variable candidate as shown in Figure 3. If 1 is used, the corresponding explanatory variable is used, and if 0, it is not used. By this coding, each individual in the genetic algorithm can be associated with one prediction model, and different explanatory variable selections can be realized depending on how 1 is selected.

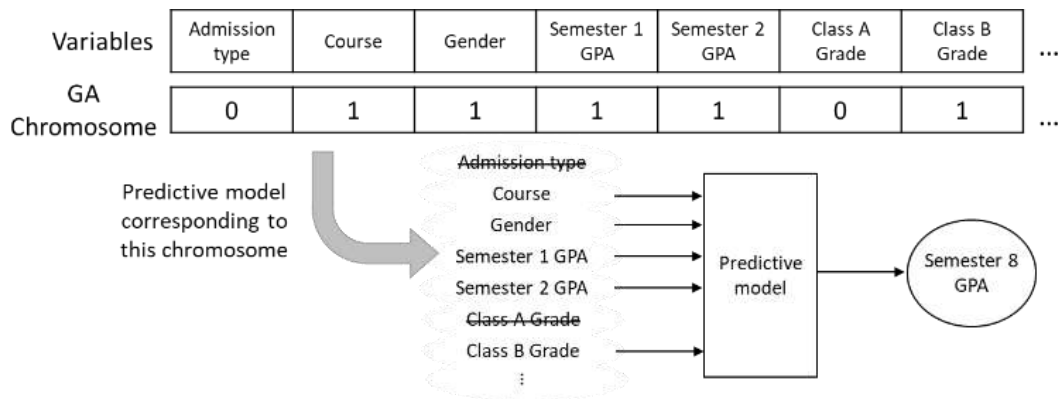


Figure 3: A chromosome expression and its relation to the predictive model.

### 3.3 Extension to Multi-objective Genetic Algorithm

The fitness of each individual in the genetic algorithm is evaluated by an arbitrarily set evaluation function. However, the evaluation criteria of the goodness of the model can be variously considered according to the purpose of use of the model. The predictive performance in AS is typical, but for other reasons, for example, the model with few explanatory variables can be said to be a good model for the reasons described in 3.2. In this way, multiple evaluation criteria should be considered in the construction of the prediction model, and this can be regarded as a multi-objective optimization problem.

Therefore, in this research, we use a multi-objective genetic algorithm, which is an extension of a genetic algorithm. As a result, it is expected that multiple model candidates can be constructed in consideration of the required trade-offs even when multiple model evaluation criteria can be considered.

The multi-objective optimization problem is stated as follows:

$$\begin{aligned} \text{Minimize} \quad & f_m(\mathbf{x}), & m = 1, \dots, M; \\ \text{subject to} \quad & g_j(\mathbf{x}) \geq 0, & j = 1, \dots, J; \\ & h_k(\mathbf{x}) = 0, & k = 1, \dots, K; \\ & x_i^{(L)} \leq x_i \leq x_i^{(U)}, & i = 1, \dots, d. \end{aligned}$$

Here,  $\mathbf{x} = (x_1, x_2, \dots, x_d)^T$  is a  $d$ -dimensional vector of the decision variable,  $x_i^{(L)}$  and  $x_i^{(U)}$  are the lower bound and the upper bound in the decision space, respectively.

In the multi-objective optimization problem, a concept of dominance is used to consider a trade-off between evaluation functions.  $x_1$  is said to dominate  $x_2$  if

$$\begin{aligned} \forall i = 1, 2, \dots, M \quad & f_i(x_1) \leq f_i(x_2) \\ \text{and} \quad \exists j = 1, 2, \dots, M \quad & f_j(x_1) < f_j(x_2). \end{aligned}$$

The solutions which are not dominated by any other solutions are called Pareto-optimal solutions. Generally, many Pareto-optimal solutions exist. A set of Pareto optimal solutions should be found because it is, in practical terms, impossible to optimize all evaluation functions at the same time in multi-objective optimization problems.

In a multi-objective genetic algorithm, a population can evolve toward a Pareto-optimal set. Since the genetic algorithm is a multipoint search algorithm, it is expected to find a Pareto-optimal set in a single simulation run.

Although various multi-objective genetic algorithms have been proposed so far, the widely known NSGA-II [14] is used in this paper. In this study, two types of fitness were assumed: predictive performance and the number of explanatory variables. The F-measure, which is generally used in classification problems, was used as the index of the predictive performances. For each individual, fivefold cross-validation was performed, and the average F-measure value was calculated.

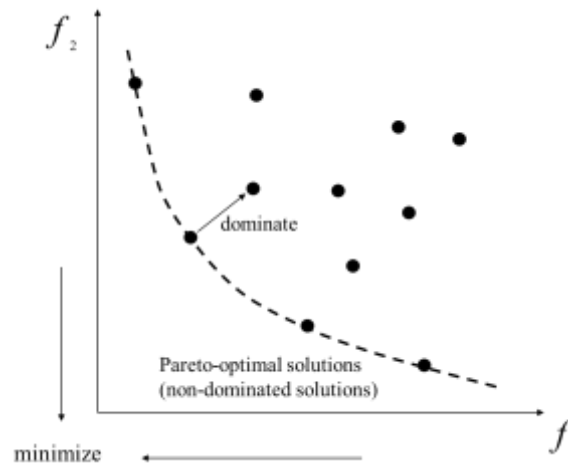


Figure 4: Pareto-optimal solutions in Multi-objective optimization.

## 4 Numerical Experiments

### 4.1 Experimental Settings

The numerical experiment was conducted using data from 181 students enrolled in university X, department Y (Economics and Business Administration) in 2015 who agreed to the use of data. One hundred and one explanatory variables were used such as admission type, courses, gender, semester GPA, acquired credits, acquisition rate of credits, sectorial GPAs, second-language tests, and class grades.

This experiment deals with the construction of AS predictive models for various AS indices. Eight academic indices will be discussed: the 8th semester cumulative GPA, and the GPAs for each of the seven areas of specialization in department Y's curriculum (Area GPA 1 to 7). Area GPA 1 represents GPA in Business Administration; Area GPA 2 represents GPA in Accounting; Area GPA 3 represents GPA in Management Science; Area GPA 4 represents GPA in Finance; Area GPA 5 represents GPA in Economics; Area GPA 6 represents GPA in History and Philosophy; and Area GPA 7 represents GPA in Mathematics and Statistics. In addition, self-evaluation at graduation of the following seven generic skills will be discussed as indices of AS: communication skills, information literacy, integrative problem thinking ability, logical thinking ability, active learning attitude, social responsibility, and cross-cultural understanding.

The predictions were based on binary classification. For the GPA indices, the training data were created with the assumption that students who had GPAs higher than the third quartile were successful. For the seven generic skill indices, we assumed that students who had answered positively (3, 4) on the self-evaluation of each skill on a four-point scale (1: strongly negative to 4: strongly positive) in the questionnaire were successful at graduation. By this definition of AS, the number of students who were successful in the experimental data was 46 (25.4%) for 8th semester cumulative GPA, 46 (25.4%) for Area GPA 1, 46 (25.4%) for Area GPA 2, 54 (29.8%) for Area GPA 3, 46 (25.4%) for Area GPA 4, 48 (26.5%) for Area GPA 5, 49 (27.1%) for Area GPA 6, 49 (27.1%) for Area GPA 7, 165 (91.2%) for communication skills, 147 (81.2%) for information literacy, 158 (87.3%) for integrated problem thinking ability, 153 (84.5%) for logical thinking ability, 164 (90.6%) for active learning attitude, 138 (76.2%) for social responsibility, and 142 (78.5%) for cross-cultural understanding.



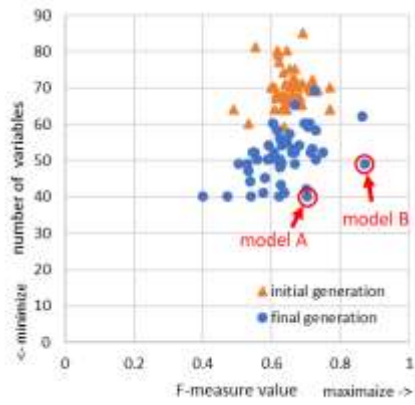
Logistic regression was used for the prediction algorithm, and the Python 3.6.4, scikit-learn [15], and DEAP [16] packages were used to implement the proposed method. The NSGA-II was employed as the multi-objective genetic algorithm, and NSGA-II was implemented using the DEAP package. The parameters of NSGA-II were 50 individuals, 500 generations, with a 0.9 (two-point crossover) crossover probability, and a 0.3 mutation probability.

## **4.2 Results**

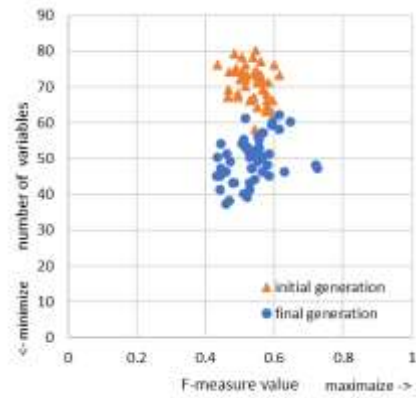
The F-measure value when all 101 variables were used is 0.674, which is used as the baseline for the following discussion. Table 1 shows the statistics of each fitness in the final generation for the experiments of each AS index. Figure 5 shows the fitness of each individual (i.e., predictive model) in the initial and final NSGA-II generations for each AS index, and Figure 6 illustrates the transition of the maximum value in the F-measure group and the minimum value in the group of the number of explanatory variables for each AS index.

Table 1: Statistics of each fitness for the final generation of NSGA-II

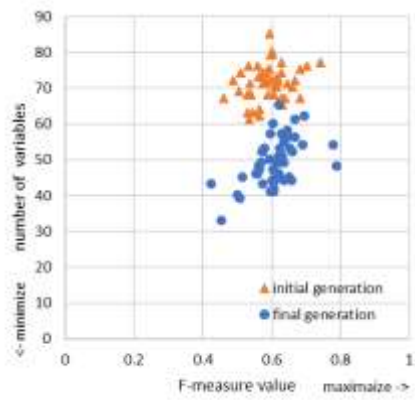
AS index	Fitness	Min	Max	Average	SD
8th semester cumulative GPA	F-measure	0.400	0.871	0.635	0.086
	# of variables	40	69	51.0	6.7
Area GPA 1	F-measure	0.432	0.725	0.536	0.065
	# of variables	37	62	49.5	6.2
Area GPA 2	F-measure	0.424	0.787	0.609	0.064
	# of variables	33	65	49.2	6.4
Area GPA 3	F-measure	0.393	0.710	0.531	0.067
	# of variables	34	66	49.5	7.1
Area GPA 4	F-measure	0.364	0.809	0.608	0.081
	# of variables	35	66	50.7	7.5
Area GPA 5	F-measure	0.419	0.804	0.620	0.079
	# of variables	34	64	51.3	7.3
Area GPA 6	F-measure	0.286	0.638	0.411	0.076
	# of variables	30	64	50.6	7.0
Area GPA 7	F-measure	0.471	0.816	0.636	0.064
	# of variables	29	64	50.7	7.2
Communication skills	F-measure	0.912	0.965	0.933	0.011
	# of variables	31	69	49.7	7.1
Information literacy	F-measure	0.810	0.897	0.845	0.018
	# of variables	39	62	60.1	6.1
Integrated problem thinking ability	F-measure	0.852	0.931	0.885	0.018
	# of variables	37	67	50.7	7.0
Logical thinking ability	F-measure	0.839	0.918	0.872	0.016
	# of variables	32	60	48.7	5.9
Active learning attitude	F-measure	0.892	0.953	0.918	0.014
	# of variables	38	60	48.1	4.8
Social responsibility	F-measure	0.759	0.855	0.793	0.021
	# of variables	38	67	50.3	6.3
Cross-cultural understanding	F-measure	0.785	0.873	0.814	0.020
	# of variables	36	62	49.6	6.4



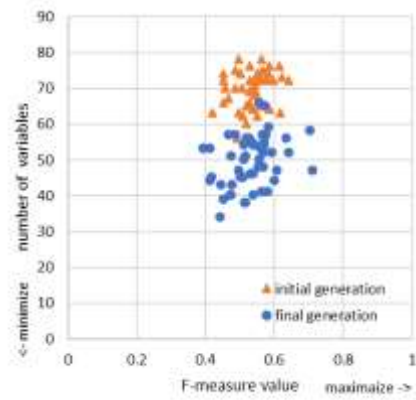
(5-a) 8th semester cumulative GPA



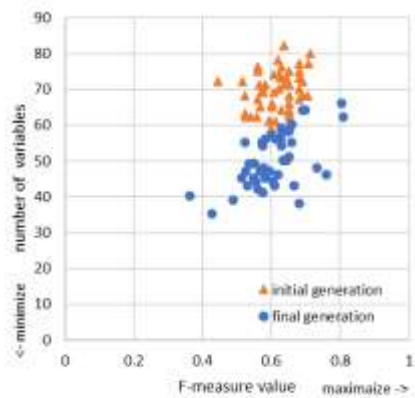
(5-b) Area GPA 1



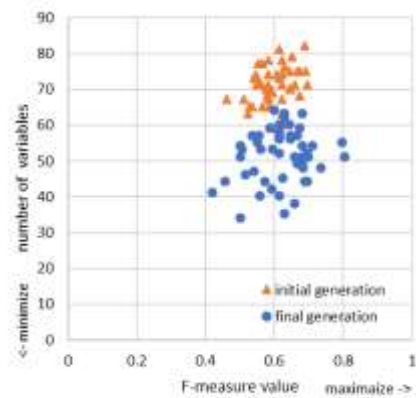
(5-c) Area GPA 2



(5-d) Area GPA 3

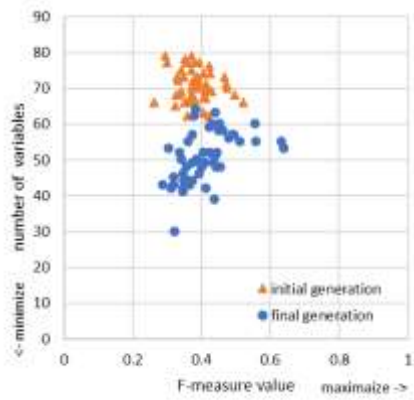


(5-e) Area GPA 4

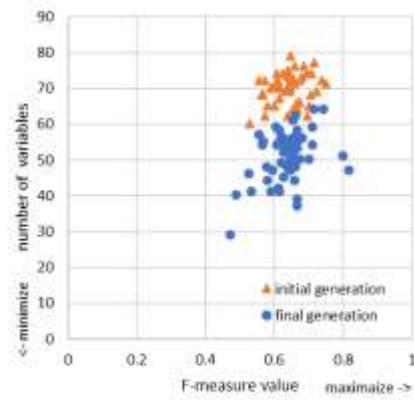


(5-f) Area GPA 5

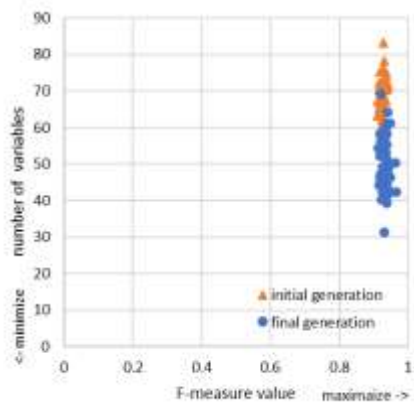
Figure 5 (a–f) : Model evaluation of each individual



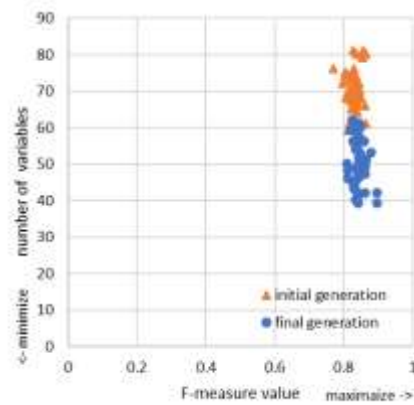
(5-g) Area GPA 6



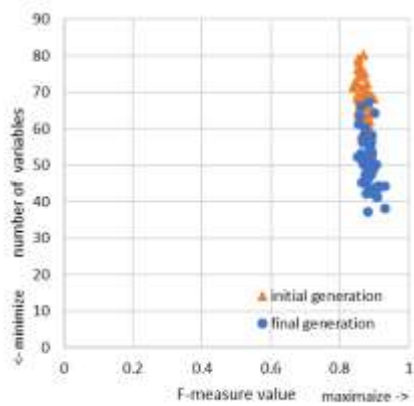
(5-h) Area GPA 7



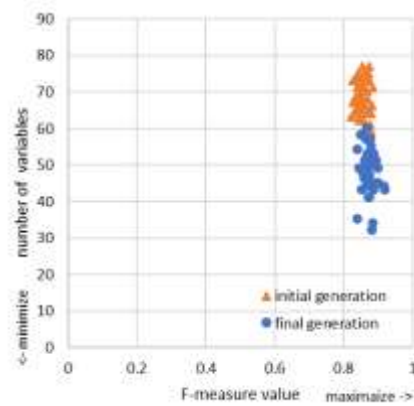
(5-i) Communication ability



(5-j) Information literacy

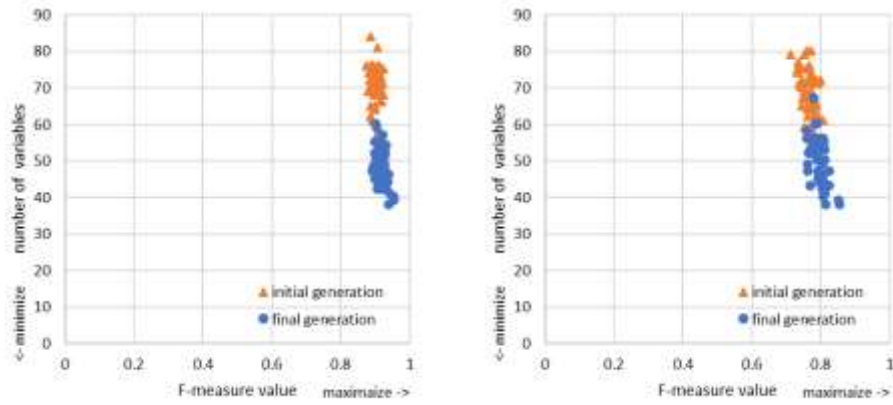


(5-k) Integrated problem thinking ability



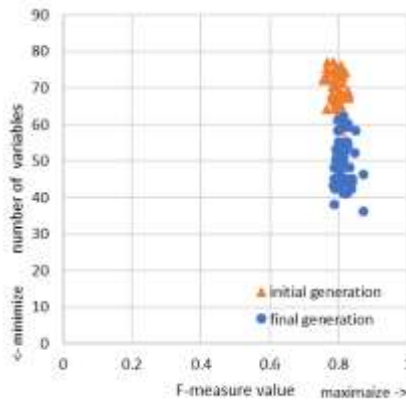
(5-l) Logical thinking ability

Figure 5 (g-l) : Model evaluation of each individual



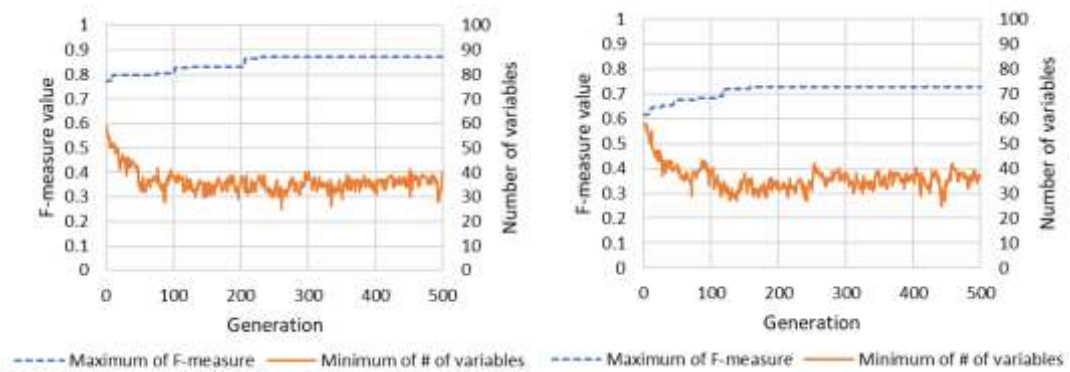
(5-m) Active learning attitude

(5-n) Social responsibility



(5-o) Cross-cultural understanding

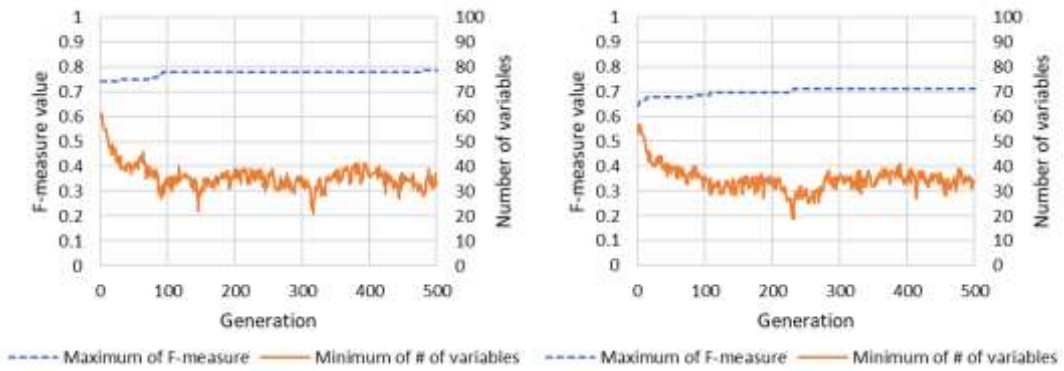
Figure 5 (m–o) : Model evaluation of each individual



(6-a) 8th semester cumulative GPA

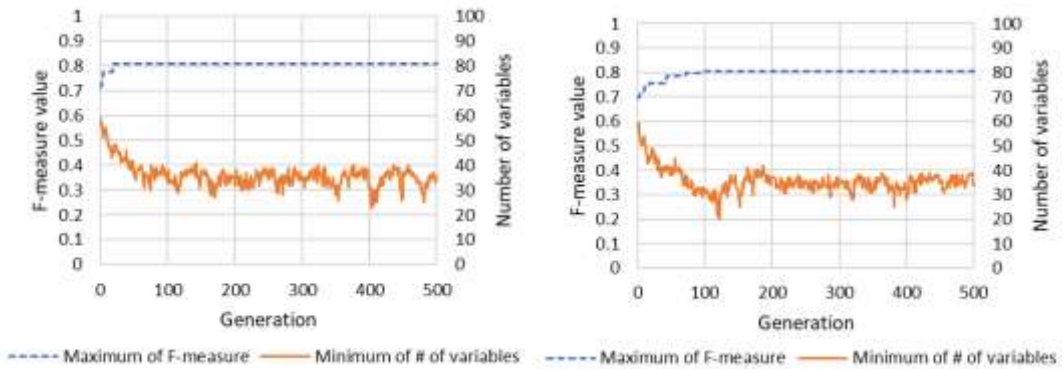
(6-b) Area GPA 1

Figure 6 (a–b): Evolution process



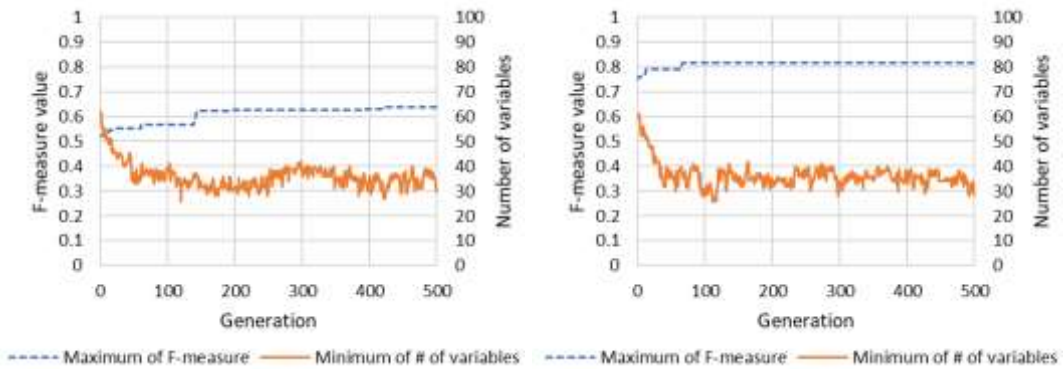
(6-c) Area GPA 2

(6-d) Area GPA 3



(6-e) Area GPA 4

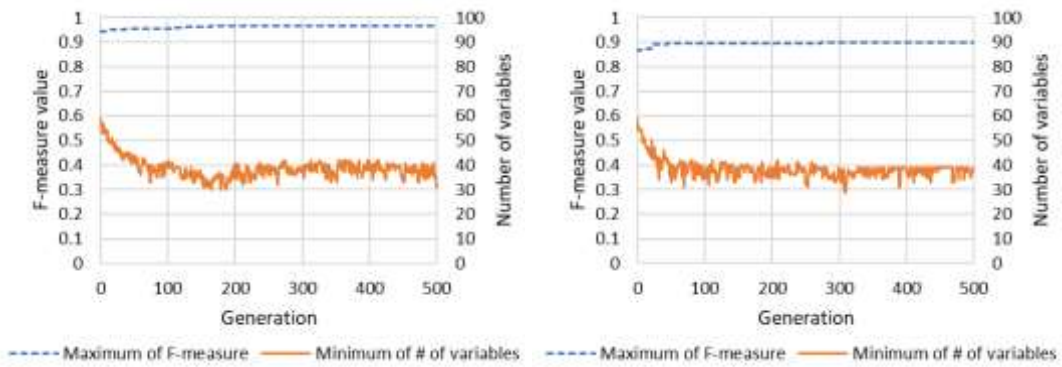
(6-f) Area GPA 5



(6-g) Area GPA 6

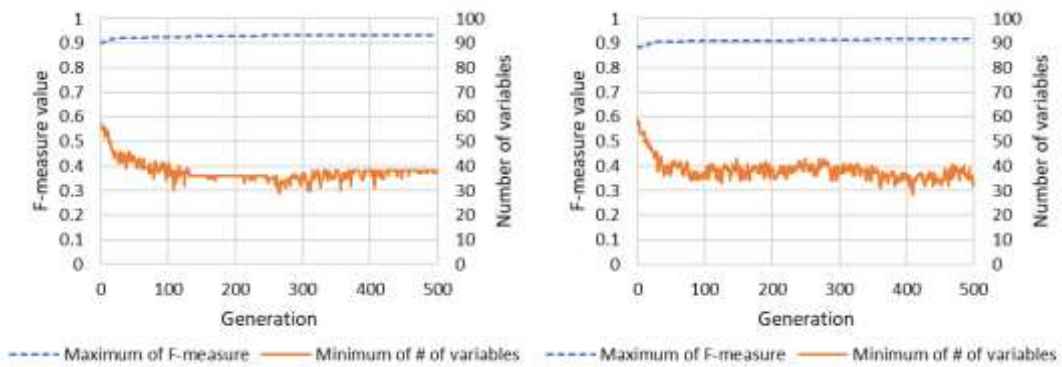
(6-h) Area GPA 7

Figure 6 (c–h): Evolution process



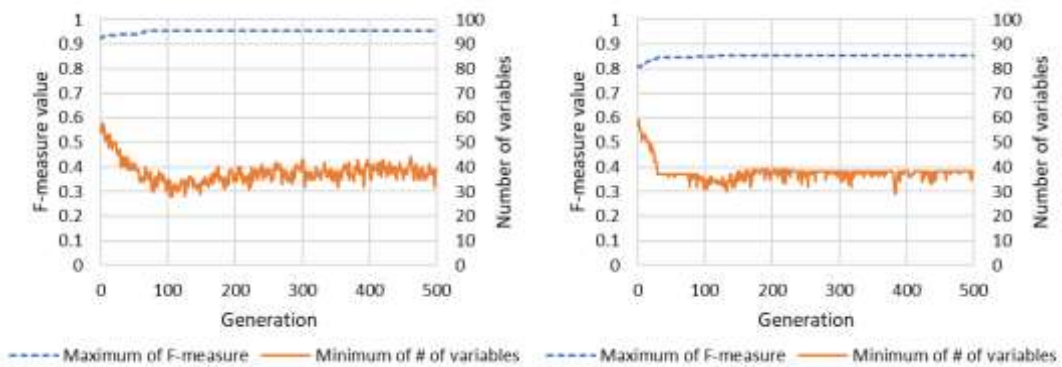
(6-i) Communication ability

(6-j) Information literacy



(6-k) Integrated problem thinking ability

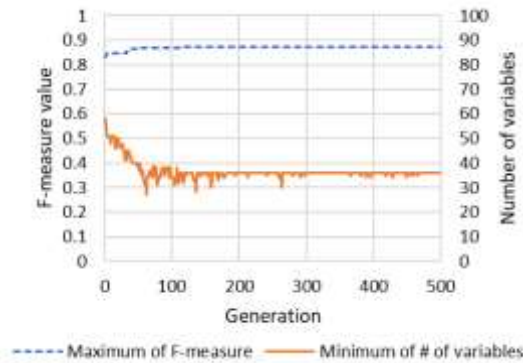
(6-l) Logical thinking ability



(6-m) Active learning attitude

(6-n) Social responsibility

Figure 6 (i–n): Evolution process



(6-o) Cross-cultural understanding

Figure 6 (o): Evolution process

From these figures, it was confirmed that as the generation progressed, a better model set was obtained while simultaneously taking into account the predictive performance and model interpretability.

Depending on the AS index, the distribution of each fitness value varies considerably. For GPA, some indices, 8th semester cumulative GPA (Figure 5-a), Area GPA 2 (Figure 5-c), Area GPA 4 (Figure 5-e), Area GPA 5 (Figure 5-f), and Area GPA 7 (Figure 5-h), have high F-measure values of up to 0.8-0.9, while in another index, Area GPA 6 (Figure 5-g) has F-measure values of up to 0.6. For the AS of generic skills, relatively high values are obtained for F-measure values, but the variance is small.

### 4.3 Discussion

From 4.2, it can be seen that the proposed method can evolve the model set in the direction that both evaluation values increase as the generation progresses, while considering the required trade-off between the prediction performance and the interpretability of the model. The decision maker can select an appropriate model from the model set of the final generation. For example, model A in the left scatterplot of Figure 5 has 40 explanatory variables and an F-measure value of 0.703, whereas model B has 49 explanatory variables and an F-measure value of 0.871, which indicates that a trade-off between performance and interpretability appears. Depending on the purpose and situation of the model's use, Model A can be selected if interpretability is more important, and Model B can be selected if prediction accuracy is more important. Since it is not always possible to decide in advance which of the two evaluation functions is more important, this method has a practical advantage insofar as it can present several possible alternatives at once while considering the two evaluation functions.

In addition to the F-measure value and the number of explanatory variables used in this study, various other evaluation functions can be set according to the requirements. For example, the classification accuracy indices, precision and recall, correspond to Type 1 and Type 2 errors, respectively, and these are in a trade-off relationship. By using the proposed method, it is possible to select an acceptable solution for the two types of statistical error under the conditions of resources and costs held by the decision maker, after seeking various alternatives. Since multi-objective genetic algorithms such as NSGA-II are applicable to multi-objective optimization with



three or more objectives, it is possible in principle to apply our method to such a wide range of evaluation criteria to be considered.

From (5-a) to (5-h) in Figure 5, we can see that the predictive model set for GPA reflects the trade-off to some extent. On the other hand, from (5-i) to (5-o) in Figure 5, we can see that the F-measure values for the self-assessment of generic skills are higher and the variances are smaller than those for the GPAs. As shown in 4.1, the ratio of students who are regarded as successful is around 25% for GPA-related AS indices and around 80-90% for generic skill-related AS indices, and the predictive performance is considered to be affected by the extent of bias between positive and negative cases of AS. In addition, since the generic skill-related indices are based on self-evaluation, the reliability of the indices needs to be examined further.

As for the evolutionary process, Figure 6 shows that the maximum F-measure value and the minimum number of variables in the population seem to improve over generations. As shown in Figure 6-e, even if the F-measure value stops improving at an early stage, it is possible to find a model with fewer variables. In this study, we determined the settings of genetic operations such as the methods of selection, crossover, and mutation, and the values of crossover and mutation probabilities based on preliminary experiments. We may be able to further improve both evaluation values by examining these settings in more detail.

## 5 Conclusions

This paper proposes a predictive model construction method that simultaneously considers the predictive performance and the interpretability of the model, taking into account the use of predictive models for advising students. The proposed method uses a multi-objective genetic algorithm to obtain a set of predictive models that reflect the required trade-off between the evaluation function for predictive performance and the desired model simplicity. The results of numerical simulations show that for AS indices such as GPA and generic skills dealt with in this paper, we can obtain a model set that reflects the trade-off while improving both evaluation function values as the generation progresses of genetic algorithm.

Future challenges include: (1) conducting numerical studies after incorporating under-sampling and other methods to deal with imbalanced data, (2) applying this method to other AS indices, and (3) conducting numerical studies when considering other evaluation functions. In addition, we plan to incorporate this method into the AS instructional support system and conduct an evaluation experiment of the system.

## Additional Statement

This is an expanded and revised version of reference [17].

## Acknowledgment

This work was supported by JSPS KAKENHI Grant Number JP17H01998 and JP19K03005.

## References

- [1] G. Siemens and R. S. J. d. Baker, “Learning Analytics and Educational Data Mining: Towards Communication and Collaboration,” *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, pp. 252-254, 2012.
- [2] A. Parnell, D. Jones, A. Wesaw, and D. C. Brooks, “Institutions’ Use of Data and Analytics for Student Success.” *EDUCAUSE: Center for Analysis and Research*, 2018.
- [3] J. Herman and M. Hilton, *Supporting Students’ College Success: The Role of Assessment of Intrapersonal and Interpersonal Competencies*. National Academies Press, 2017.
- [4] G. A. Rice and A. B. Russell, “Refocusing Student Success: Toward a Comprehensive Model,” in *The Handbook of Institutional Research*, R. D. Howard, G. W. McLaughlin, and W. E. Knight, Eds. John Wiley & Sons, 2012, pp. 237–255.
- [5] Central Council for Education, “Gakushi Katei Kyoiku No Kochiku Ni Mukete (Toshin),” Ministry of Education, Culture, Sports, Science and Technology, 2008. (in Japanese)
- [6] C. Brooks and C. Thompson, “Predictive Modelling in Teaching and Learning,” in *Handbook of Learning Analytics*, 1st ed., Alberta, Canada: Society for Learning Analytics Research, 2017, pp. 61–68.
- [7] Y. Hayashi, Y. Watanabe, H. Matsukawa, T. Matsuda, M. Tsubakimoto, S. Tateishi, and H. Yamashita, “Development of SVM based Risk Detector for Retention of University Students,” *Proceedings of the 15th Annual Hawaii International Conference on Education*, p. 31, January 2017.
- [8] O. Nelles, *Nonlinear System Identification*, Springer, 2001.
- [9] K. E. Arnold and M. D. Pistilli, “Course signals at Purdue: using learning analytics to increase student success,” *Proceedings of 2nd International Conference on Learning Analytics and Knowledge*, pp. 267–270, April 2012.
- [10] S. M. Jayaprakash, E. W. Moody, E. J. M. Lauria, J. R. Regan, and J. D. Baron, “Early Alert of Academically At-Risk Students: An Open Source Analytics Initiative,” *Journal of Learning Analytics*, vol. 1, no. 1, pp. 6–47, May 2014.
- [11] K. P. Murphy, *Machine Learning: a Probabilistic Perspective*, MIT Press, 2012.
- [12] T. Bäck, *Evolutionary Computation*, Oxford University Press, 1996.
- [13] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, 1st ed., USA: Addison-Wesley Publishing Company, 1989.
- [14] K. Deb, *Multi-Objective Optimization using Evolutionary Algorithms*, New York, USA: John Wiley & Sons, 2001.
- [15] scikit-learn, “Machine Learning in Python,” 2021. [online]. Available: <https://scikit-learn.org/stable/>. [Accessed: March. 31, 2021].

- [16] DEAP Project, “DEAP documentation,” 2021. [online], Available: <https://deap.readthedocs.io/en/master/>. [Accessed: March. 31, 2021].
- [17] N. Kondo, T. Matsuda, Y. Hayashi, H. Matsukawa, M. Tsubakimoto, Y. Watanabe, S. Tateishi, and H. Yamashita, “Academic Success Prediction based on Important Student Data Selected via Multi-objective Evolutionary Computation,” Proceedings of the 9th International Congress on Advanced Applied Informatics (IIAI-AAI 2020), 2020.