

Estimation of Students' Learning States using Bayesian Networks and Log Data of Learning Management System

Nobuhiko Kondo ^{*}, Toshiharu Hatanaka [†]

Abstract

Over the last decade, learning analytics (LA) and its related fields have rapidly developed. LA has become a robust concept of Institutional Research (IR) or student support. In this study, building a model of learning process by a Bayesian network from a large amount of educational data has been studied. Moreover, we propose the Bayesian network model for the learning process based on the log data accumulated from the learning management system (LMS). From the numerical simulation results, the proposed approach can be used to predict certain learning states of students.

Keywords: Learning analytics, Institutional research, Bayesian networks, Learning management system, Learning states.

1 Introduction

Over the last decade, big data have emerged with special applications in numerous areas, including marketing research, health care services, business, social media, and drug development. With some statistical techniques, big data can help making an adequate decision, finding useful patterns or rule set, and detecting a change, fault, or chance. For example, to perform successful decision-making, the big data which are gathered or recorded momentarily should be analyzed by some methods such as analytics and modeling, which have an important role in big data usage.

In the learning sector, educational institutions possess the vast amount of student data, which include demographic data, learning results or performance, registration data, and access log of learning management system (LMS). These data have become be able to be collected in real-time by rapid development of information technology. Recently, learning analytics (LA) and educational data mining have particularly received much attention [1] and are tackling the problems occurring on educational institutions by methods such as data mining, machine learning, and big data analysis.

Conversely, several educational institutions, particularly colleges and universities, need to maintain high retention rates; therefore, enrollment management is important. Academic analytics, an application of analytics in the level of institution or politics, has also been

^{*} Tokyo Metropolitan University, Tokyo, Japan

[†] Osaka University, Osaka, Japan

proposed [2] and is highly concerned in Institutional Research (IR). Recent studies such as early detection of at-risk students with LA has been considered herein to detect students prone to high dropout risk early and intervene them effectively, by using the educational big data [3].

This study aims to develop a practical method to model students' learning states with a large amount of educational data. Application of this method requires a suitable mathematical model. Herein, a graphical probabilistic model is employed because its interpretation is easy for users. Furthermore, for the utilization of the model by IR, we consider constructing the model on the transition of the learning states using a Bayesian network from a large amount of log data accumulated in the LMS and examining its availability.

The remainder of this paper is organized as follows. In Section 2, we review analytics, such as LA and educational data mining, in higher education. In Section 3, we introduce a method using a Bayesian network for modeling students' learning states via LMS log data. In Section 4, the results of the numerical simulation are presented. The conclusions are provided in Section 5.

2 Analytics in Higher Education

Recent rapid development of information communication technology has facilitated innovation in several areas. For example, EdTech represents a movement that aims to increase innovation in education using the appropriate technology [4], and educational big data have received much attention to achieve this objective. For effective application of the educational big data where applying analytics is important, technologies such as machine learning, artificial intelligence, data mining, and statistical learning have been introduced.

Since 2000, big data analytics in education has been studied actively; particularly, LA is one of the major research areas in education. On Call for Papers of the first International Conference on Learning Analytics and Knowledge (LAK), LA was defined as follows [5]: *Learning analytics is the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs*. Educational data mining is also the major research field of educational analytics, focusing on how to apply techniques of data mining in education. A concept of academic analytics had also been proposed in the late 2000s with reference to analytics in the level of institution or politics [2]. In academic analytics, quality assurance in education or the accountability of university is the main focus. Thus, the targets are learning outcomes or retention of students and it has a close relation with the viewpoint of enrollment management in IR.

Over the past decade, LA has been applied in the organizational settings such as Course Signal at Purdue University [6], predictive modeling in Marist College Open Academic Analytics Initiative [3], Competency Map in Capella University [7], Expected Graduation Rate Calculator in UCLA Higher Education Research Institute [8], and Grade Craft in the University of Michigan [9].

In IR and academic analytics, approaches for predicting the learning state based on educational data have attracted increasing attention. Moreover, several studies have used machine learning or data mining techniques for the prediction of learning outcomes or the early detection of at-risk students [10][11][12][13][14][15]. In these studies, the problems are assumed to be pattern classification problems, where the educational data available at a certain point are considered as explanatory variables, whereas the future states of the

students are considered as objective variables.

3 Bayesian Networks about Learning Process Constructed with LMS Log Data

Many studies on analytics in higher education have been attempted to assist students according to their situation adaptively; thus, their learning states have to be appropriately modeled. Although several data mining or machine learning techniques are available, understanding their inner process and meaning is still important for users in practical situation. Therefore, the interpretable model should be used for modeling students' learning states. In this study, a Bayesian network is used for such modeling.

3.1 Bayesian Network

A Bayesian network is a graphical model for a discrete probabilistic distribution [16]. A Bayesian network represents some conditional dependency of the random variables by the directed acyclic graph and conditional probability tables. Fig. 1 shows an example of a Bayesian network structure. The nodes represented by the circles reveal the random variables and the directed links show the dependency of the random variables. The network structure and the conditional probability table are estimated on the basis of given data. Fig. 2 shows an example of the conditional probability table. Fig. 1 and Fig. 2 are illustrative examples, and a_i , b_i , and e_i indicate abstract labels that can be taken by the nodes A, B, and E, respectively. Prior knowledge can be used to determine the network structure.

When some variables, called evidence, are observed, the other variables can be inferred as the posterior distribution of variables. This process is called probabilistic inference. According to [17], in a simple case where there is one parent node X_1 and one child node X_3 for a certain node X_2 , and e^+ is the parent node observation information and e^- is the child node observation information, the posterior probability of X_2 when the evidence e is obtained, $P(X_2|e)$, is as follows using Bayes' theorem:

$$\begin{aligned}
 P(X_2|e) &= P(X_2|e^+, e^-) \\
 &= \frac{P(e^-|X_2, e^+)P(X_2|e^+)}{P(e^-|e^+)}
 \end{aligned}
 \tag{1}$$

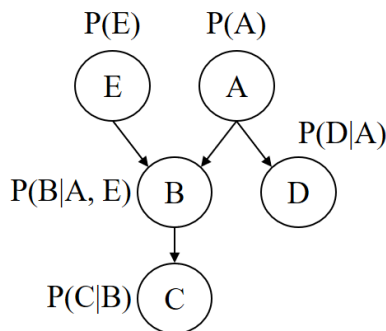


Figure 1: An example of a Bayesian network structure.

Node	Condition		Conditional Probability	
	A	E	$B = b_1$	$B = b_2$
B	a_1	e_1	0.5	0.5
	a_1	e_2	0.2	0.8
	a_2	e_1	0.3	0.7
	a_2	e_2	0.1	0.9

Figure 2: An example of the conditional probability table.

Here, because e^+ and e^- is conditionally independent when X_2 is fixed, let $\alpha = 1/P(e^-|e^+)$ be a normalization constant, and let the probability of propagation from the child node be $P(e^-|X_2, e^+) = P(e^-|X_2) = \lambda(X_2)$. In addition, let the probability of propagation from the parent node be $P(X_2|e^+) = \pi(X_2)$, Using that e^- is independent of the value X_2 , P can be transformed as follows:

$$P(X_2|e) = \alpha\lambda(X_2)\pi(X_2) \quad (2)$$

$$\begin{aligned} \lambda(X_2) &= \sum_{X_3} P(e^-|X_2, X_3)P(X_3|X_2) \\ &= \sum_{X_3} P(e^-|X_3)P(X_3|X_2) \end{aligned} \quad (3)$$

$$\pi(X_2) = \sum_{X_1} P(X_2|X_1)P(X_1|e^+) \quad (4)$$

By applying these formulas recursively to upstream of the parent node and downstream of the child node based on appropriate prior probabilities for the terminal nodes, observations, and conditional probability tables, $P(X_2|e)$ can be calculated. Even when there are a plurality of parent nodes and child nodes, this may be extended using a conditionally independence property. Depending on whether or not the network structure has a loop, an efficient calculation method such as a junction tree algorithm or a Loopy BP method is known. There are many documents that can be referred to regarding Bayesian networks and their probabilistic inference methods. For more details, refer to [17], for example.

Since the Bayesian network has a high capability of representing the probabilistic structure of an intended system, it can serve as a useful method for modeling uncertainty. Therefore, it has been applied in several domains such as fault diagnosis, pattern recognition, and data mining. Human modeling [18] [19] is one of the successful fields of a Bayesian network.

3.2 Modeling of the Processes of Learners with Bayesian Networks

The modeling of the processes of learners has been studied considerably for the application in intelligent tutoring systems or adaptive e-learning systems [20]. A Bayesian network or a Bayesian approach is often used for student modeling because of its availability. In an example, a Bayesian network model is built into the test theory because of its capability to estimate the degree of knowledge and understand learners [21][22]. In another example, a system construct a decision tree model or a Bayesian network model with nodes can be used to extract from a learning record data accumulated in an LMS, and the system generates adaptive messages based on the prediction of future grades on a certain e-learning course [23][24][25]. In another study, a learning style is estimated using a Bayesian network based on the action recorded in an e-learning course, and the learner is supported adaptively based on the estimated learning style [26]. Millan et al. summarized the application of a Bayesian network for student modeling from basic concept to actual examples [27]. These examples are about a learner's learning process of a certain course, which reveal a student modeling at more micro level than that with the modeling of the holistic learning process and the educational institution outcomes.

Conversely, a Bayesian network is often widely used for the analysis and prediction in education and learning areas, not only in such e-learning courses [28]. In an example,

a Bayesian network is used for the analysis of the effect of education on student category [29]. In the study, a model is constructed by classifying students into a category based on the results of a questionnaire survey, and this method can be used to analyze the educational effectiveness and career levels of many courses. However, it is uncertain whether the results of the analysis offer feedback to the students, which could be used to predict the growth of each student category. In another study, a Bayesian network is used for recommending learning methods suited to students' characteristics by the predicting their grades [30]. In the study, the Bayesian network model is constructed based on the results of a questionnaire survey of junior high school students that focuses on learning methods and grades. Another study uses a Bayesian network to predict future grades based on the attendance data and grades of the first year by assuming student support provided at the educational institution level [31].

3.3 Bayesian Network about Students' Learning Process

Using a Bayesian network, a framework for modeling the transition process of undergraduate students' learning states was proposed [32]. In this framework, a Bayesian network about learning process is constructed using the following process:

1. Some appropriate variables are selected as candidates of the nodes from the variables of the baccalaureate degree program.
2. The network structure is determined while considering the chronological sequence.
3. The conditional probability table is learned from training data.

This Bayesian network can be used in the following cases:

- early detection of students with high dropout risk using the probabilistic inference,
- estimation of future learning states using the probabilistic inference with the data up to an arbitrary point in the baccalaureate degree program,
- learning support based on the results of the probabilistic inference to help take certain actions in the future.

In the previous studies, only offline student data such as demographic attributes, attendance in classes, and grades at the end of the term, have been dealt with. In this study, we consider using the online LMS log data to construct a Bayesian network about the learning process and investigate the further possibility of this model.

4 Numerical Experiment

4.1 Experimental Settings

In the numerical experiments, LMS log data at University X are used. This LMS is operated on the whole university. Students in this university are mandated to use the LMS, both in classes and for browsing important information such as announcements from the university. Hence, it is expected to reflect a level of commitment for learning in the university. In this LMS, whenever some operation is performed, a log is recorded with a set of student ID, operation content and operation time as one record as shown in Table 1.

In the experiments, log data of the first semester of 548 students admitted in 2015 were used. Because the operation contents were diverse, a part of the log that reflected the learning states was extracted. In particular, log records about four different actions were extracted:

- logging-in,
- browsing of an announcement from the university,
- starting an assignment of two compulsory subjects,
- submission of the assignment.

As students are required to submit the assignments of the two compulsory subjects each time using the LMS, these logs are selected to reflect the learning states. The log data of the 15 weeks of the first semester were used. Therefore, the number of extracted records was 146,723.

Although it seems that a better prediction can be made if there is data such as scoring results of the assignment, such data does not exist in the LMS log used. Therefore, in this study, we investigated how far we can predict from the features that can be extracted from the available data.

Table 1: Sample of LMS log data.

ID	Date	Operation	Category	Detail
AAAAAA	20**-04-01 12:00:00.472812	Login	Home	
BBBBBB	20**-04-01 12:00:14.121184	Login	Home	
CCCCCC	20**-04-01 12:01:02.395736	Login	Home	
AAAAAA	20**-04-01 12:01:23.024680	Switching function	Information	
BBBBBB	20**-04-01 12:02:00.193742	Switching function	Settings	
AAAAAA	20**-04-01 12:02:00.936451	Lesson start	Report	Class A
⋮	⋮	⋮	⋮	⋮

Table 2: Features generated from log data (When “*” at the end of the feature name is “1,” it indicates the first to fifth weeks, “2” indicates the sixth to tenth weeks, and “3” indicates the eleventh to fifteenth weeks.)

Features	Contents
<i>Login*</i>	Total number of logins at a certain 5 weeks
<i>Info*</i>	Total number of browsing notifications for a certain 5 weeks
<i>Start*</i>	Total number of starting assignments in a certain 5 weeks
<i>Submit*</i>	Total number of assignment submissions in a certain 5 weeks

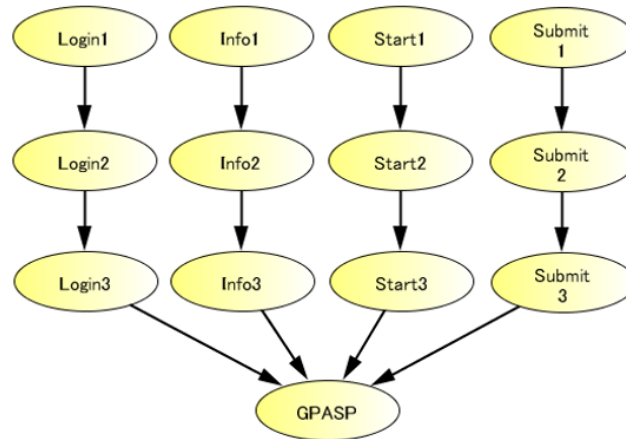


Figure 3: Structure of the Bayesian network named BN_{simple} .

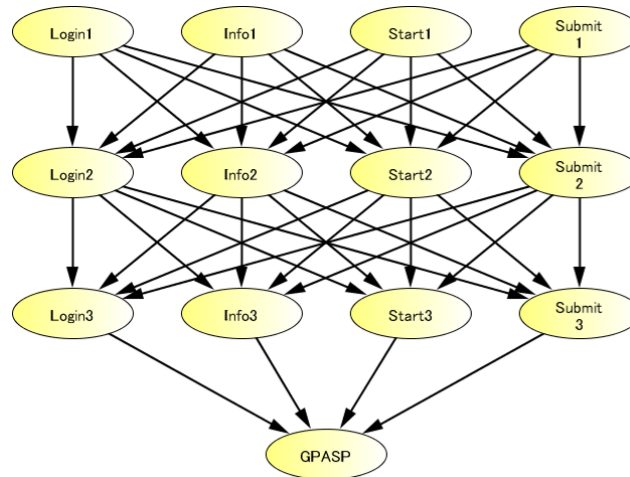


Figure 4: Structure of the Bayesian network named $BN_{complex}$.

Total 12 features were generated and used for the probability variable (node) to construct a Bayesian network. A list of the features is shown in Table 2. The values of the total number of corresponding log records are calculated every five weeks for above-mentioned four actions. Given these 12 features and Grade Point Average (GPA) of the first semester as nodes, we developed two types of Bayesian network structure. A simple one is named BN_{simple} and a complex one is named $BN_{complex}$. These network structures are shown in Fig. 3 and Fig. 4. Each structure was determined before considering the chronological sequence.

Each variable was discretized by the k -means method, and GPA was discretized into three states, and other variables into four states. Therefore, GPA was divided into the range of $[0, 1.6]$, $(1.6, 2.9]$, and $(2.9, 4]$, but we call them *Low*, *Mid*, and *High*, respectively, for convenience. Sample sizes of each class are as follows: *Low* is 119 (21.7%), *Mid* is 252 (46.0%) and *High* is 177 (32.3%).

As GPA is discretized into three classes, the Bayesian network outputs the probability

value for each class. In the experiments, *Low* is used as the prediction label when the probability value of the *Low* class reaches a threshold T_l or more. When the probability value of the *Low* class does not reach T_l and the probability value of the *High* class reaches a threshold T_h or more, *High* is used as the label. Otherwise *Mid* is used as the label.

BayoLink version 7.1.0 developed by the National Institute of Advanced Industrial Science and Technology was used for the experiments.

4.2 Results (1) - Performance Prediction

To confirm the performance of the model, we examined some experiments about the prediction of the learning states. Two types of experiments about the prediction of GPA are implemented as follows:

1. The prediction is performed based on the log data up to the fifth week, i.e., only with four variables of the first to fifth weeks: *Login1*, *Info1*, *Start1*, and *Submit1*.
2. The prediction is performed based on the log data up to the tenth week, i.e., with eight variables of the first to tenth weeks: *Login1*, *Info1*, *Start1*, *Submit1*, *Login2*, *Info2*, *Start2*, and *Submit2*.

The performance of the predictive ability is evaluated using five-fold cross-validation. Cross validation is usually used to learn a model with high generalization ability. In this paper, the sample size is not large, and the result of cross validation is used as a performance evaluation of the proposed method. Precision and recall are used for the performance indicator. Let True Positive (*TP*) be the number of positive instances of the predicted positive, False Positive (*FP*) be the number of negative instances of the predicted positive, and False Negative (*FN*) be the number of positive instances of the predicted negative; precision is calculated as $TP/(TP + FP)$ and recall is calculated as $TP/(TP + FN)$. Precision is an indicator of the correctness of the output of the predictive model and recall is an indicator of the ability to detect the relevant instances in a dataset.

In the following experiments, the threshold T_h varied from 0.3 to 0.45 in the step size 0.05, and T_l varied from 0.2 to 0.4 in the step size 0.05. Experimental results for BN_{simple} are shown in Fig. 5 to Fig. 12, and the results for $BN_{complex}$ are shown in Fig. 13 to Fig. 20. There seems no significant difference between the performance of BN_{simple} and $BN_{complex}$.

The threshold value takes the tradeoff between precision and recall. Using the same network structure and the span of log data, there is a tendency that the larger the recall value, the lower the precision value. In addition, there is a tendency for the precision and recall values being higher than logs accumulate, which indicates that the prediction becomes easier as logs accumulate. The experimental result indicates that GPA can be predicted to some extent by only actions such as *Login**, *Info**, *Start**, and *Submit**, not by learning results.

The settings of T_l affect the performances of *Low* class and *Mid* class. Similarly, the settings of T_h affect the performances of *High* class and *Mid* class.

In future, we intend to establish an approach to set the threshold values appropriately. When the IR method is considered, the importance of precision and recall values will depend on the university situation. Thus, a method to obtain multiple candidates for the model based on the trade-off instead of obtaining only one model is required.

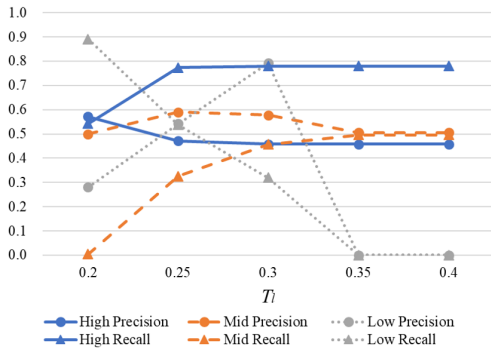


Figure 5: Prediction performance of $BN_{simple}(T_h = 0.3)$ with log data up to the fifth week.

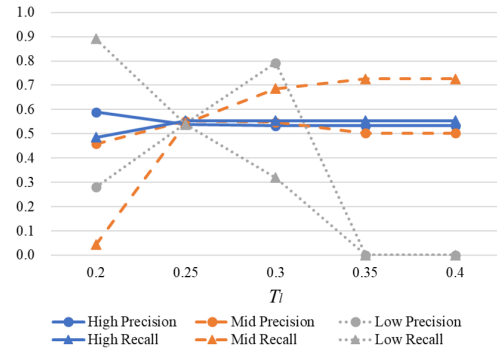


Figure 6: Prediction performance of $BN_{simple}(T_h = 0.35)$ with log data up to the fifth week.

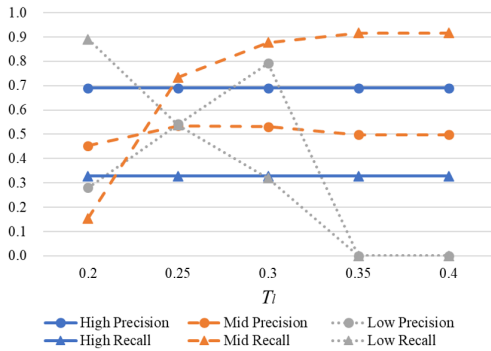


Figure 7: Prediction performance of $BN_{simple}(T_h = 0.4)$ with log data up to the fifth week.

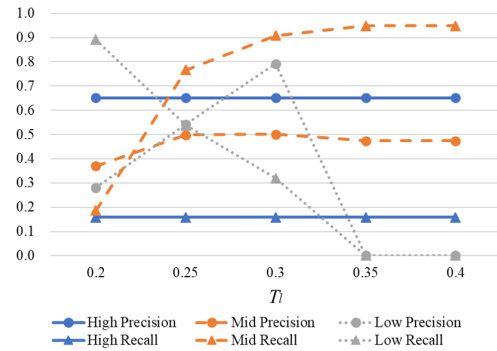


Figure 8: Prediction performance of $BN_{simple}(T_h = 0.45)$ with log data up to the fifth week.

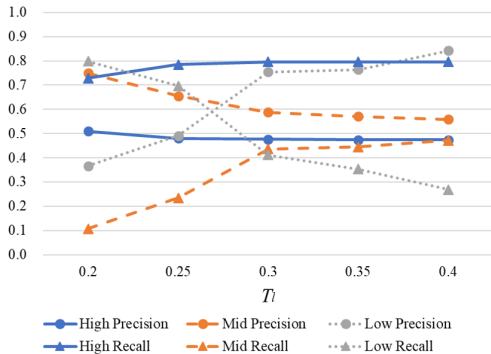


Figure 9: Prediction performance of $BN_{simple}(T_h = 0.3)$ with log data up to the tenth week.

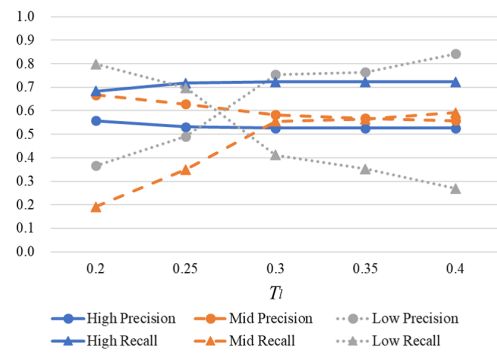


Figure 10: Prediction performance of $BN_{simple}(T_h = 0.35)$ with log data up to the tenth week.

4.3 Results (2) - Probabilistic Inference

The proposed model can be used for the learning support based on the probabilistic inference results for taking certain actions in the future. In this experiment, initially

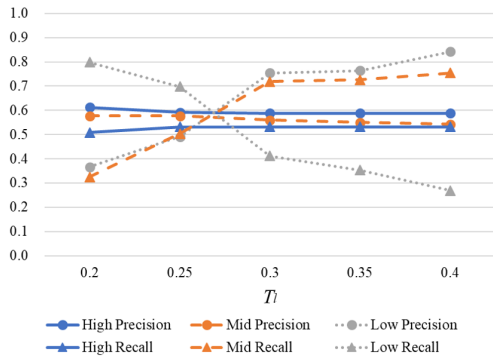


Figure 11: Prediction performance of $BN_{simple}(T_h = 0.4)$ with log data up to the tenth week.

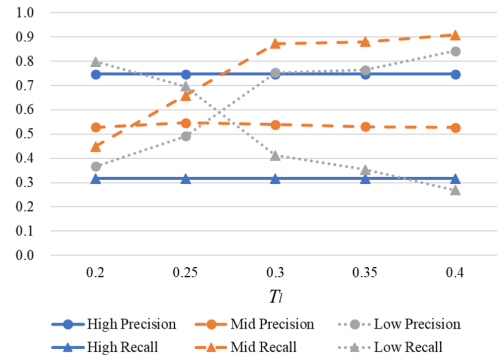


Figure 12: Prediction performance of $BN_{simple}(T_h = 0.45)$ with log data up to the tenth week.

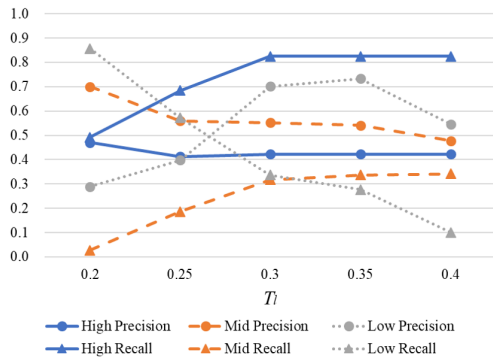


Figure 13: Prediction performance of $BN_{complex}(T_h = 0.3)$ with log data up to the fifth week.

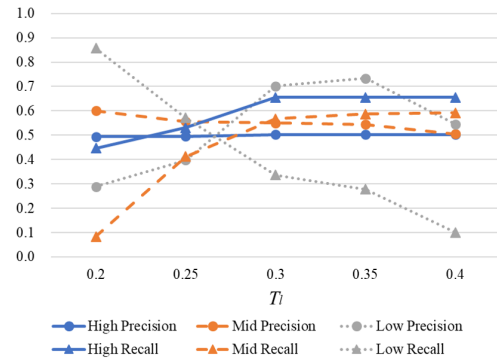


Figure 14: Prediction performance of $BN_{complex}(T_h = 0.35)$ with log data up to the fifth week.

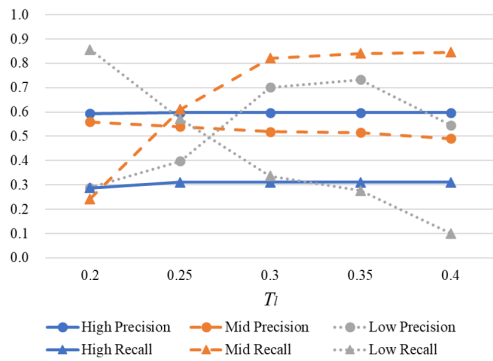


Figure 15: Prediction performance of $BN_{complex}(T_h = 0.4)$ with log data up to the fifth week.

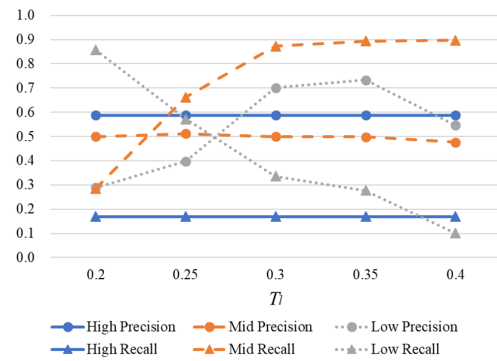


Figure 16: Prediction performance of $BN_{complex}(T_h = 0.45)$ with log data up to the fifth week.

a Bayesian network with the $BN_{complex}$ structure is learned with the training data; then certain values of evidence are set to some nodes step by step. We set the scenario in which a student does not

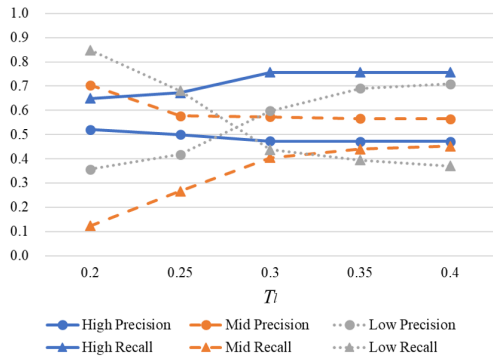


Figure 17: Prediction performance of $BN_{complex}(T_h = 0.3)$ with log data up to the tenth week.

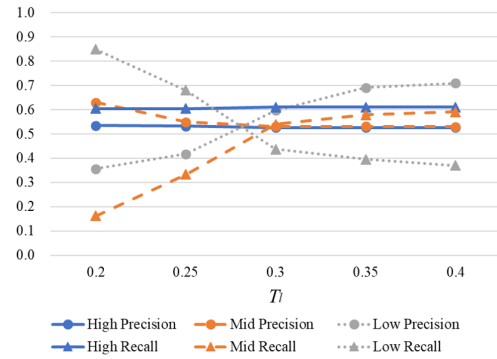


Figure 18: Prediction performance of $BN_{complex}(T_h = 0.35)$ with log data up to the tenth week.

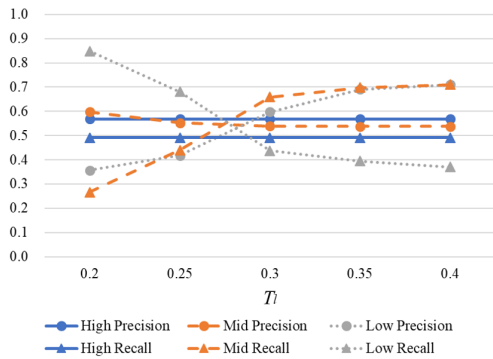


Figure 19: Prediction performance of $BN_{complex}(T_h = 0.4)$ with log data up to the tenth week.

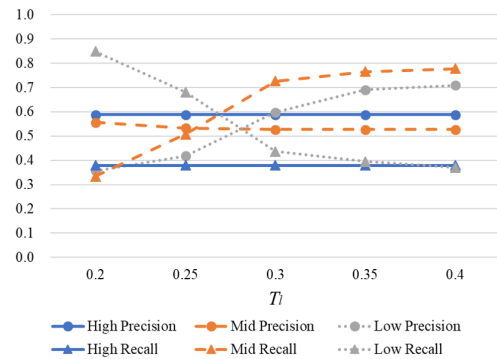


Figure 20: Prediction performance of $BN_{complex}(T_h = 0.45)$ with log data up to the tenth week.

use the LMS actively for the first five weeks, then he/she uses it actively used for the next five weeks.

The inferred probability values of each node with evidence based on the scenario are shown in Fig. 21 and Fig. 22. Fig. 21 indicates that inactive students for the first five weeks continue to be inactive later, thus, they are likely to get a lower grade. Fig. 22 indicates that such student's learning state may become active and get a higher grade if active action is taken for the next five weeks. Such information can help students or their supporters, e.g., teacher, and learning support staff.

5 Conclusions

In this study, we proposed an approach to estimate students' learning states using a Bayesian network based on LMS log data and examined its applicability using actual data. The proposed approach can predict certain learning states of students. In future, we plan to develop models with a higher temporal resolution to tackle the determination problem of the network structure, and consider the application for student support.

Login1		Info1		Start1		Submit1	
State	Value	State	Value	State	Value	State	Value
[0, 38]	1.0000	[0, 6]	1.0000	[0, 19]	1.0000	[0, 7]	1.0000
(38, 71]	0.0000	(6, 19]	0.0000	(19, 28]	0.0000	(7, 11]	0.0000
(71, 140]	0.0000	(19, 41]	0.0000	(28, 42]	0.0000	(11, 12]	0.0000
(140, ∞)	0.0000	(41, ∞)	0.0000	(42, ∞)	0.0000	(12, ∞)	0.0000

Login2		Info2		Start2		Submit2	
State	Value	State	Value	State	Value	State	Value
[0, 26]	0.8788	[0, 4]	0.8788	[0, 16]	0.5455	[0, 8]	0.6970
(26, 46]	0.0909	(4, 14]	0.1212	(16, 28]	0.2727	(8, 14]	0.2424
(46, 80]	0.0303	(14, 30]	0.0000	(28, 47]	0.0909	(14, 16]	0.0303
(80, ∞)	0.0000	(30, ∞)	0.0000	(47, ∞)	0.0909	(16, ∞)	0.0303

Login3		Info3		Start3		Submit3	
State	Value	State	Value	State	Value	State	Value
[0, 24]	0.7147	[0, 7]	0.7309	[0, 13]	0.3535	[0, 8]	0.4689
(24, 41]	0.1794	(7, 20]	0.1513	(13, 26]	0.4271	(8, 15]	0.3223
(41, 74]	0.0655	(20, 41]	0.0769	(26, 37]	0.1073	(15, 17]	0.0868
(74, ∞)	0.0404	(41, ∞)	0.0409	(37, ∞)	0.1121	(17, ∞)	0.1220

GPA_SP	
State	Value
high	0.1759
mid	0.4258
low	0.3984

Figure 21: Probability values of each node with the evidences up to the fifth week.

Login1		Info1		Start1		Submit1	
State	Value	State	Value	State	Value	State	Value
[0, 38]	1.0000	[0, 6]	1.0000	[0, 19]	1.0000	[0, 7]	1.0000
(38, 71]	0.0000	(6, 19]	0.0000	(19, 28]	0.0000	(7, 11]	0.0000
(71, 140]	0.0000	(19, 41]	0.0000	(28, 42]	0.0000	(11, 12]	0.0000
(140, ∞)	0.0000	(41, ∞)	0.0000	(42, ∞)	0.0000	(12, ∞)	0.0000

Login2		Info2		Start2		Submit2	
State	Value	State	Value	State	Value	State	Value
[0, 26]	0.0000	[0, 4]	0.0000	[0, 16]	0.0000	[0, 8]	0.0000
(26, 46]	0.0000	(4, 14]	0.0000	(16, 28]	0.0000	(8, 14]	0.0000
(46, 80]	1.0000	(14, 30]	1.0000	(28, 47]	1.0000	(14, 16]	1.0000
(80, ∞)	0.0000	(30, ∞)	0.0000	(47, ∞)	0.0000	(16, ∞)	0.0000

Login3		Info3		Start3		Submit3	
State	Value	State	Value	State	Value	State	Value
[0, 24]	0.0000	[0, 7]	0.0000	[0, 13]	0.0000	[0, 8]	0.0000
(24, 41]	0.4000	(7, 20]	0.6000	(13, 26]	0.6000	(8, 15]	1.0000
(41, 74]	0.2000	(20, 41]	0.4000	(26, 37]	0.4000	(15, 17]	0.0000
(74, ∞)	0.4000	(41, ∞)	0.0000	(37, ∞)	0.0000	(17, ∞)	0.0000

GPA_SP	
State	Value
high	0.4931
mid	0.3811
low	0.1258

Figure 22: Probability values of each node with the evidences up to the tenth week.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number JP16K16331, JP16H03082 and JP17H01998.

References

- [1] G. Siemens and R. S. J. d. Baker, "Learning Analytics and Educational Data Mining: Towards Communication and Collaboration," in *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, pp. 252–254, 2012.
- [2] J. P. Campbell and D. G. Oblinger, "Academic Analytics," *EDUCAUSE Review*, 2007.
- [3] 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, 2014.
- [4] CDW, EdTech Magazine, 2018. [Online]. Available: <http://www.edtechmagazine.com/>. [Accessed: Feb. 28, 2019].
- [5] The 1st International Conference on Learning Analytics and Knowledge, Call For Papers, July 22, 2010. [Online]. Available: <https://tekri.athabascau.ca/analytics/call-papers>. [Accessed: Feb. 28, 2019].
- [6] K. E. Arnold, and M. D. Pistilli, "Course signals at Purdue: using learning analytics to increase student success," in *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, pp. 267–270, 2012.
- [7] J. Grann, "LAK15 Case Study 3: FlexPath: Building Competency-based, Direct Assessment Offerings," *Learning Analytics Review*, no. LAK15-3, 2015.
- [8] Higher Education Research Institute, HERI Expected Graduation Rate Calculator, 2018. [Online]. Available: <https://heri.ucla.edu/GradRateCalculator.php>. [Accessed: Feb. 28, 2019].
- [9] The University of Michigan, GradeCraft, 2018. [Online]. Available: <https://www.gradecraft.com/>. [Accessed: Feb. 28, 2019].
- [10] S. Amenomori, T. Matsuda, and T. Mori, "One Strategy of Institutional Research for Faculty Development: Focused on the Area of Teaching and Learning," *Kyoto University Researches in Higher Education*, vol. 18, pp. 1–10, 2012. (in Japanese)
- [11] A. Ohtomo, Y. Iwayama, and T. Mohri, "On-Campus Data Utilization: Working on IR (Institutional Research) in Universities," *FUJITSU*, vol. 65, no. 3, pp. 41–47, 2014.(in Japanese)
- [12] A. Djulovic and D. Li, "Towards Freshman Retention Prediction: A Comparative," *International Journal of Information and Education Technology*, vol. 3, no. 5, pp. 494-500, 2013.

- [13] N. Kondo and T. Hatanaka, "Modelling of Students' Learning States Using Big Data of Students through the Baccalaureate Degree Program," *Transactions of Japanese Society for Information and Systems in Education*, vol. 33, no. 2, pp. 94–103, 2016. (in Japanese)
- [14] H. Lakkaraju, E. Aguiar, C. Shan, D. Miller, N. Bhanpuri, R. Ghani, and K. L. Addison, "A Machine Learning Framework to Identify Students at Risk of Adverse Academic Outcomes," in *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1909–1918, 2015.
- [15] H. Fujiwara, "Gakugyo O Chudan Suru Gakusei No Yosoku Model Kochiku Ni Tsuite," *Journal of Institutional Research and University Education*, vol. 5, pp. 8–22, 2016. (in Japanese)
- [16] G. F. Cooper and E. Herskovits, "A Bayesian Method for the Induction of Probabilistic Networks from Data," *Machine Learning*, vol. 9, no. 4, pp. 309–347, 1992.
- [17] Y. Motomura and H. Iwasaki, *Bayesian Network Gijutsu*, Tokyo Denki University Press, 2006. (in Japanese)
- [18] E. Horvitz, J. Breese, D. Heckerman, D. Hovel, and K. Rommelse, "The Lumiere Project: Bayesian User Modeling for Inferring the Goals and Needs of Software Users," in *14th National Conference on Uncertainty in Artificial Intelligence*, pp. 256–265, 1998.
- [19] C. Conati, A. S. Gertner, K. VanLehn, and M. J. Druzdzel, "On-Line Student Modeling for Coached Problem Solving Using Bayesian Networks," in *Proceedings of the 6th International Conference on User Modeling*, pp. 231–242, 1997.
- [20] K. Chrysafiadi and M. Virvou, "Student modeling approaches: A literature review for the last decade," *Expert Systems with Applications. An International Journal archive*, vol. 40, no. 11, pp. 4715–4729, 2013.
- [21] M. Ueno, H. Ohnishi, and K. Shigemasu, "Proposal of a Test Theory with Probabilistic Network," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences (Japanese Edition)*, vol. J77-A, no. 10, pp. 1398–1408, 1994 (in Japanese).
- [22] M. Ueno, H. Ohnishi, and K. Shigemasu, "Adaptive Testing Based upon Teacher Domain Knowledge Structures," *Japan journal of educational technology*, vol. 18, no. 1, pp.15–24, 1994 (in Japanese).
- [23] M. Ueno, "Animated Pedagogical Agent based on Decision Tree for e-Learning," in *Proceedings of the 5th IEEE International Conference on Advanced Learning Technologies*, pp. 188–192, 2005.
- [24] M. Ueno and T. Okamoto, "Bayesian Agent in e-Learning," in *Proceedings of the 7th IEEE International Conference on Advanced Learning Technologies*, pp. 282–284, 2007.
- [25] M. Ueno. "Intelligent LMS with an Agent that Learns from Log Data," *The Journal of Information and Systems in Education*, vol. 7, no. 1, pp. 3–14, 2008.

- [26] S. Schiaffino, P. Garcia, and A. Amandi, "eTeacher: Providing personalized assistance to e-learning students," *Computers & Education*, vol. 51, no. 4, pp.1744–1754, 2008.
- [27] E. Millan, T. Loboda, and J. L. P. de la Cruz, "Bayesian networks for student model engineering," *Computers & Education*, vol. 55, no. 4, pp.1663–1683, 2010.
- [28] R. G. Almond, R. J. Mislevy, L. S. Steinberg, D. Yan, and D. M. Williamson, *Bayesian Networks in Educational Assessment*, Springer, 2015.
- [29] M. Tsubaki, and A. Iwasaki, "Verification of measurement and prediction precision on an analysis method of the educational effectiveness according to the student types using a Bayesian Network," *Kyoiku Joho Kenkyu*, vol. 26, no. 4, pp. 25–36, 2011 (in Japanese).
- [30] H. Maruyama, and T. Ohtsuka, "Recommendation on Styles of Studying for Junior High School Students Using Bayesian Networks," *Transactions of the Institute of Systems, Control and Information Engineers*, vol. 29, no. 11, pp. 476–485, 2016 (in Japanese).
- [31] H. Itoh, K. Itoh, and K. Funahashi, "Forecasting Students' Grades Using Bayesian Network Models and an Evaluation of Their Usefulness," *The Journal of Information and Systems in Education*, vol. 11, no.1, pp. 32–41, 2012.
- [32] N. Kondo and T. Hatanaka, "Modeling of Learning Process based on Bayesian Network," *Japan Journal of Educational Technology*, vol. 41, no. 3, pp. 271–281, 2017 (in Japanese).
- [33] NTT Data Mathematical Systems Inc., BayoLink. [Online]. Available: <http://www.msi.co.jp/bayolink/>. [Accessed Feb. 28, 2019].