Tracing Knowledge States through Student Assessment in a Blended Learning Environment

Indriana Hidayah* and Ebedia Hilda Am

Department of Electrical and Information Engineering, Universitas Gadjah Mada
Jl. Grafika No. 2, Yogyakarta, 55281, Indonesia
**Corresponding author. Email: indriana.h@ugm.ac.id*

Abstract— Blended learning has recently acquired popularity in a variety of educational settings. This approach has the advantage of being able to autonomously monitor students' knowledge states using the collected learning data. Moodle is the most widely used learning management system in blended learning environments. Students can access Moodle to obtain supplementary materials, exercises, and assessments to complement their face-to-face meetings. However, its performance can be improved by more effectively tailoring students' skills and pace of learning. Several studies have been conducted on knowledge tracing; however, we have not discovered any studies that particularly investigate knowledge tracing in a blended learning setting with Moodle as a component. This study proposes a scheme for assessment using the features of the Moodle quiz platform. The assessment data is used to conduct knowledge tracing with the Bayesian Knowledge Tracing (BKT) model, which improves interpretability. The aforementioned data were collected from information engineering undergraduate students who completed 88 exercises that assessed 23 knowledge components within the course. We measure RMSE and MAE to evaluate the performance of the BKT model on our dataset. Furthermore, we compare the knowledge tracing performance to other well-known datasets. Our results show that the BKT model performed better with our dataset, with an RMSE of 0.314 and an MAE of 0.197. Moreover, the BKT model can be used to assess student performance and determine the level of mastery for each knowledge component. Thus, the outcomes can be applied to personalized learning in the future.

Keywords— bayesian knowledge tracing; blended learning; knowledge state; Moodle

I. INTRODUCTION

Blended learning, a combination of face-to-face meetings and online learning, is being widely used in the post-pandemic period. This is due to the many benefits offered by the method. It is retaining the values of face-to-face meetings while at the same time gaining the advantages of online learning which is conducted by using computer-based learning systems. Studies show that many advantages are obtained with the implementation of blended learning, such as flexibility [1], efficiency, and effectiveness in teaching and learning activities [2], [3]. Other studies also reveal that blended learning, when applied together with a student-centered instructional approach, has the potential to increase students' engagement, motivation, and self-direction [4], [5]. Furthermore, blended learning can provide valuable digital data that can be utilized for learning analytics [6], [7].

There are various tasks in learning analytics; one of them is knowledge tracing. It is used to monitor the development of students' knowledge and/or skills during a learning process. Monitoring student knowledge states is crucial to facilitate the personalization of teaching materials [8], interventions or treatments [9], instructions, and exercises [10]. In other words, understanding the student's current knowledge state is a prerequisite to being able to adapt the learning environment according to students' skills and abilities [11].

In a face-to-face class, knowledge tracing can be done manually by teachers by directly observing students' activity, engagement, and performance. Nevertheless, for a large class with many students, this job becomes very challenging. Therefore, data collected from online sessions of a blended learning method can be utilized to perform knowledge tracing automatically, which allows for a personalized learning experience that is believed to be more effective.

Several approaches to knowledge tracing are available such as probabilistic and deep learning. The advantage of the probabilistic approach is that it can trace the student's knowledge state by modeling skill as a binary latent variable (unlearned state and learned state) and then calculating the transitions between the two variables for each answer [12]–[14]. However, the probabilistic approach still has limitations in terms of accuracy. Another approach for knowledge tracing is deep learning. This approach tends to have better accuracy, consequently, it gains popularity in the last five years [11]. Specifically, the method is called as deep knowledge tracing (DKT). Despite the fact that DKT offers advantage in term of accuracy, it lacks interpretability [15]–[17]. A deep learning approach does not seem to be an ideal solution, especially when explicit underlying theory and interpretability are important [18]. Interpretability is required to track changes in the student's knowledge state so that an estimate of the student's cognitive state as well as the skills that have been successfully mastered can be found. Meanwhile, this study has proven that probabilistic approach extensions using Bayesian Knowledge Tracing (BKT) can have predictive performance similar to a deep learning approach and can maintain interpretability, including psychologically meaningful model parameters [19], [20].

Previous researchers have conducted several studies [13], [9], [21] related to BKT. The first study proposed the BKT model by extending learning conditions from two (unlearned and learned) to three (unlearned, learning, and learned) [13].

Received 29 September 2023, Revised 26 January 2024, Accepted 29 January 2024.
DOI: https://doi.org/10.15294/jte.v15i2.47861

The datasets utilized for this study were the Assistments Math datasets for the years 2004–2005, 2005–2006, and 2006–2007, which were acquired from WPI-Assistments via DataShop. The RMSE value derived from the three datasets, specifically when considering the combination of knowledge components (KCs), falls within the range of 0.475. Meanwhile, the following study developed the BKT model by taking into account the relationship between knowledge [9]. This research uses datasets from the online education system Bridge to Algebra. The dataset is a set of steps from each topic completed by the student and systematically recorded. In this context, the BKT here serves to calculate the probability of a student having mastered certain knowledge. Once the student exceeds the minimum probability threshold, the system will cease providing exercises related to that knowledge. Approximately 0.395 is the RMSE value derived from all the skills. The other study estimated a student's state of knowledge using the concepts of partially correct and partially known [21]. This study used six well-known datasets, such as Algebra I 2005–2006, Algebra I 2006–2007, Bridge to Algebra 2006–2007, ASSISTments 2009–2010, ASSISTments 2012–2013, and ASSISTments 2017. Particularly in the ASSISTments 2012–2013 dataset, the model generated an RMSE value of approximately 0.05. According to the findings of these studies, BKT accuracy may be enhanced. This method promises the needed performance for automatic tracking of student knowledge in a blended learning environment due to BKT's superiority in terms of interpretability.

Datasets from prior studies are gathered from intelligent tutoring systems. In those systems, assessment schemes are built and implemented specifically for knowledge tracing purpose. Meanwhile, many educational institutions implement blended learning methods by utilizing learning management system (LMS) platforms which are configured based on the need of the institutions. Moodle is one of many commonly used LMSs. Therefore, it is important to design, implement, and evaluate knowledge tracing scheme in Moodle.

Referring to the previous explanation, this research is motivated by two problems. First, we were unable to locate any prior studies that employed BKT for knowledge tracing in a blended learning environment with Moodle serving as the online learning platform. Due to Moodle's widespread use as an LMS, it is crucial to investigate its integration with BKT. In the meantime, certain intelligent tutoring systems are used for the majority of knowledge tracing systems. Secondly, the efficacy of BKT in prior studies can be enhanced through a reduction in the RMSE value.

The purpose of this study was to track the academic performance and cognitive states of students by analyzing a dataset obtained from exercises and quizzes provided through Moodle. The efficacy of knowledge tracing is also assessed through comparisons with other datasets. In order to accomplish this, we propose a scheme for preparing assessments using Moodle that will facilitate the knowledge tracing process, which is detailed in Section II. The knowledge tracing was then performed, and the findings were presented in the results and discussion Section III.

II. Method

The study was conducted in an actual blended learning environment at the School of Information Engineering, Faculty of Engineering, Universitas Gadjah Mada, Indonesia. The course on Algorithm and Data Structure was conducted using Moodle as a learning management platform for the online component of a blended learning environment. Students are required to master 23 knowledge components (KCs) through a series of exercises. The 23 KCs are decomposed from learning outcomes which are defined by the extended syllabus written in the curricula. Knowledge tracing is used to monitor the students' level of mastery at each KC.

In this section, we explain the setup of the blended learning environment and assessment by using Moodle as the source of the dataset. After that, we also described the collected dataset as well as the knowledge tracing method.

A. Blended Learning Setup in Moodle

Prior to delving into the knowledge tracing method, it is crucial to take into account several types of blended learning implementation, as they will determine the effectiveness of the teaching method. Blended learning implementation typically involves two main categories of instructional design: teacher-centered and student-centered approaches [5]. Within a teacher-centered approach, the utilization of the online learning system gives rise to three different configurations. A teacher-centered approach with screen configuration utilizes an online learning system to facilitate face-to-face meetings by providing students with online materials, exercises, and assessments. Besides, the teacher-centered approach also has a scene and cockpit configuration, whereas the student-centric approach has a crew, metro, and ecosystem configuration [5].

Screen configuration in a blended learning environment is implemented in this study. The online learning system is utilized for providing various learning materials such as slides, videos, links, and e-books, as well as conducting formative and summative assessments. Moodle is one of the most popular open-source LMS that may be utilized for the development, management, and distribution of learning materials. Moodle with hypermedia resources facilitates planning, supervision, control, personal reflection on students' own practice, and real-time feedback upon task completion [17]. This strategy thus promotes students' involvement in their own learning process. Moodle has been shown to be more effective and efficient than traditional learning environments in attaining in-depth and high-quality learning if it includes hypermedia resources that can be adapted to the learning pace of each student through a personalized learning environment [17], [18]. Furthermore, Moodle provides useful features to be explored to conduct assessments and collect data for knowledge tracing.

B. Assessment Setup in Moodle

Each blended learning environment has five pedagogical dimensions: 1) combination, which combines online and face-to-face learning strategies; 2) mediatization, which includes e-learning and instructional design; 3) mediation, which deals with media effects and behaviour; 4) teacher and student mentoring, which includes cognitive, metacognitive, and motivational components; and 5) degree of openness or flexibility, which lets students choose the learning methods and resources they want to use [5]. This study uses the first pedagogical dimension, which involves a combination of face-to-face learning with educators and Moodle as an online learning platform. Moodle was configured to accomplish the study's objectives. First, students enrol in the data structure course available on Moodle. After that, students can access the materials that educators have uploaded to Moodle and participate in classroom sessions with educators. Then, when the educator has finished delivering the material, the student will be asked to work on the quiz using Moodle.

Figure 1 shows the scheme for the preparation of students' assessments. We started from a defined extended syllabus document of Algorithm and Data Structure course which is used in the institution. Based on the document, a set of learning outcomes for the course is given. We decomposed each learning outcome into a set of inter-related KCs. Sets of KCs are outlined because they are vital for knowledge tracing. It enables the process of tracing student knowledge states for each KC. The KCs are also used as a guide for items or questions generation step. For each KC we develop several items. These items or questions which are generated are then stored in the question bank which is available in Moodle. After that, the quiz is setup in Moodle and items for the quiz are imported from the question bank.

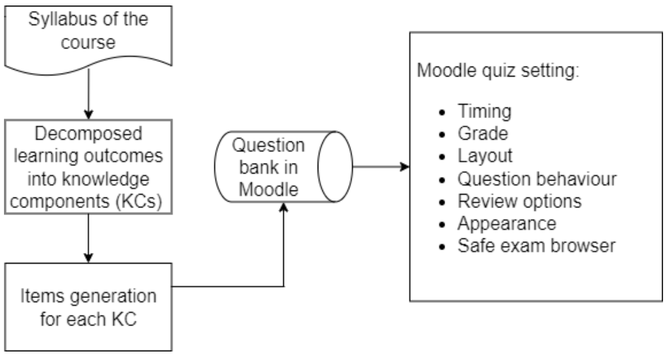

Figure 1. Scheme for assessment setup using Moodle

TABLE I. QUIZ SETTINGS ON MOODLE

No	Settings	Objective
1	Open the quiz	The quiz's opening date and time settings
2	Close the quiz	The quiz's closing date and time settings
3	Time limit	A time limit has been set for each attempt
4	Grade to pass	Minimum grade requirements for students
5	Attempts allowed	The maximum number of trials a student is allowed to do
6	Grading method	The highest score from all the trials conducted is utilized by the evaluation system
7	Shuffle within questions	Each student works on the questions in a distinct order, as they are displayed at random
8	Review options	Students will only receive information regarding their scores and the number of completed trials while the quiz is still available

It has been mentioned that this study is conducted on an Algorithm and Data Structure course. There are three learning outcomes specified in the extended syllabus document. For this study, we developed the assessment plan for two learning outcomes as follows.

- Students can explain and implement Abstract Data Type (ADT) i.e., tree and graph data structures.
- Students can explain and implement various data sorting algorithms and search algorithms.

The 23 KCs consist of tree structure, characteristics of trees, types of trees, application of each tree type, implementation of trees using C++, tree operations (tree traversal), tree operations (binary search trees), graph structure, types of graphs, graph terminologies, graph implementation, graph traversal, linear search (operating principle), linear search (performance analysis), binary search (operating principle), binary search (performance analysis), linear and binary search, insertion sort (operating principle), insertion sort (performance analysis), selection sort (operating principle), selection sort (performance analysis), bubble sort (operating principle), bubble sort (performance analysis). For each KC, we create many questions or items that utilize different question formats offered by Moodle, including multiple-choice and true/false.

The assessment's applicable settings and their objectives are described in Table I. The quiz can be customized with several settings, such as time limitations, a minimum passing grade requirement, the number of allowed attempts, question behavior, and review options. In the assessment settings, we enforce a requirement that students must get a minimum score of 85 within a maximum of three attempts.

C. Dataset

This study employs quizzes and question data obtained from the Algorithm and Data Structure course on Moodle, along with student data retrieved from the academic information system. The data are then preprocessed to generate the Moodle 2023 dataset. Datasets for tracing knowledge states are collected in unique learning environments; thus, each dataset has different characteristics. The dataset used in this study is collected in a realistic learning environment, which enhances its applicability. It consists of 32,820 records, which were gathered from 252 students who enrolled in the Algorithm and Data Structure class, as shown in Table II. Furthermore, the dataset captures students progression from one knowledge component to the next, which is crucial for understanding students knowledge or skill from time to time.

In addition, the performance of BKT is evaluated using the Bridge to Algebra 2006-2007 and ASSISTments 2009-2010 datasets. Bridge to Algebra 2006–2007 is a dataset of the Intelligent Tutoring System (ITS) known as Cognitive Tutor. This dataset comprises a comprehensive record of all stages of each mathematical topic, encompassing 12 KCs. Subsequently, the ITS at Worcester Polytechnic Institute provided publicly the ASSISTments 2009–2010 dataset. This dataset comprises 111 KCs related to mathematics. Details regarding each of the datasets included in this study are given in Table II, including the total number of students, records, exercises, and KCs.

TABLE II. STATISTICS OF ALL DATASET

Statistics	Datasets		
	Bridge to Algebra 2006–2007	**ASSISTments 2009–2010**	**Moodle 2023**
Students	587	4,217	252
Records	16,857	525,534	32,820
Exercises	550	26,688	88
Knowledge concepts	12	111	23

D. Knowledge Tracing Process

Figure 2 illustrates the process of tracking the knowledge states of students. The process commences with the acquisition of data from Moodle and academic information systems. Moodle serves as a repository of information pertaining to quizzes, questions, and students' scores, while the academic information system functions as the repository of students' personal information. Following that, the data from both sources undergoes the data preprocessing stage, encompassing several steps such as data cleaning, data transformation, data integration, and data reduction, resulting in the development of the Moodle 2023 dataset. Afterwards, the Moodle 2023 dataset proceeds to the student modeling stage, where student activity records are utilized to estimate students' knowledge states. During this stage, the BKT model is used to measure student performance and cognitive state. After that, in order to assess the effectiveness of the BKT model, this study used the root mean squared error (RMSE) and mean absolute error (MAE).

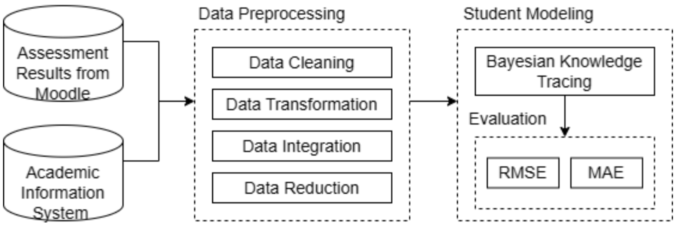

Figure 2. Process of tracing student knowledge states

1) Data Preprocessing

Various data preprocessing tasks were conducted, comprising data cleaning, data transformation, data integration, and data reduction. Data cleaning followed the extraction of data from both Moodle and the academic information system. After that, the student's grade on each attempt was individually scrutinized; if the student had achieved a passing grade on the preceding attempts, the subsequent attempt was eliminated. The data then undergoes data transformation, which involves converting wide data into long data. The following step involves data integration, wherein the data from quizzes, questions, and students is merged into a unified table. Next, the data reduction steps involve eliminating irrelevant features, such as surname, first name, email address, state, started on, completed, and grade, as they are unnecessary for the student modeling. Table III contains a detailed summary of the features obtained during the data preprocessing stage.

TABLE III. DESCRIPTION OF FEATURES IN THE MOODLE 2023 DATASET

Feature	Description
user_id	Student ID at the time of the quiz
problem_id	ID of the problem to be addressed
problem_name	Name of the problem to be addressed
attempt_count	Number of trials (number of times students submitted answers)
answer_type	Type of question containing multiple-choice and true-false options
skill_id	ID of the skill for each problem
skill_name	Name of the skill in each problem
student_class_id	Class ID consists of classes A, B, and C
school_id	ID of the study program pursued by the student
time_taken	The number of seconds it takes a student to complete a quiz from start to finish
correct	The student answers comprise 0 (incorrect responses) and 1 (correct responses)

2) Student Modeling

BKT is a two-state learning model that transitions between unlearned and learned states using a rule that disregards forgetting and the absence of the opposite transition rule [11], [22]. The possibility of something occurring in the unlearned state is that the student guesses correctly [23]. In contrast, the only possibility in the learned state is that the student will make a mistake. Both conditions are modeled as binary variables, 0 for the unlearned state and 1 for the learned state [24].

BKT consists of four parameters: P(L), P(T), P(G), and P(S) [9]. Parameters P(L) and P(T) are learning parameters used to indicate knowledge states. P(L) represents the students' knowledge state, while P(T) denotes the probability of transitioning from an unlearned state to a learned state. P(L) = 0 indicates that the student has not yet mastered KC, whereas P(L) = 1 indicates the inverse. Meanwhile, if the value of the P(T) is 0, there is no transition from the unlearned to the learned state. The other two parameters, P(G) and P(S), are performance parameters. P(G) represents the probability that a student responded successfully at the time of the unlearned state, while P(S) denotes the probability that a student failed to answer at the time. When P(G) is 0, the student answered incorrectly during the time they did not master KC, and when P(S) is 0, the student answered correctly during the time they mastered KC. Therefore, when both P (G) and P (S) are 0, the student's response reflects their knowledge mastery.

The unlearned state is represented by (1-P(L)), and the learned state is represented by (P(L)). BKT is binary-based student modeling, so there are only two kinds of answers: incorrect and correct. When the answer is incorrect, there are two possibilities: the student fails to answer in the unlearned state (1-P(G)) or makes a mistake in the learned state (P(S)). When the answer is correct, the probability is that the student is able to answer in the unlearned state (P(G)) or not make a mistake in the learned state (1-P(S)).

BKT calculates the probabilities of the KC learned in the experiment ($_t$), with the learning parameters P(L_t) and t > 0, so that P(L_t) is the probability that the KC has been understood after the experiment. P(L_t) is to be updated after an attempt using (1)-(3) [13]. Correct$_t$ means that the performance of the *t*-th chance to apply KC is true, and Incorrect$_t$ means the opposite. Evidence can be Correct$_t$ or Incorrect$_t$, which means the performance of the *t*-th chance of applying KC.

$$P(L_{t-1}|Correct_t) = \frac{P(L_{t-1})(1-P(S))}{P(L_{t-1})(1-P(S)) + (1-P(L_{t-1}))P(G)} \quad (1)$$

$$P(L_{t-1}|Incorrect_t) = \frac{P(L_{t-1})P(S)}{P(L_{t-1})P(S) + (1-P(L_{t-1}))(1-P(G))} \quad (2)$$

$$\mathrm{P}(L_t) = P(L_{t-1}|evidence_t) + \left(1 - P(L_{t-1}|evidence_t)\right) * P(T) \quad (3)$$

Figure 3 illustrates how the two phases of estimating student knowledge state and performance can represent the BKT model representation of the preceding equation. In Figure 3(a), the black circles represent the KC's learning state at time t (knowledge states), while the white circles represent the student's performance at time t. The arrow between the black circles is the probability of the knowledge state transitioning from time *t-₁* to time *t*. Whereas the arrow between the black circles and the white circles is the probability of the performance transition at that time. P(L_0) and P(L_{t-1}) mark students mastering KC at time steps 0 and t. Meanwhile, figure 3(b) shows the relationship between bisection knowledge states (unlearned and learned states) and bisection performance (incorrect and correct). P(T) is the probability that the learning state will transition from unlearned to learned following the application of a KC, whereas P(S) and P(G) represent the probabilities of slipping and guessing, respectively.

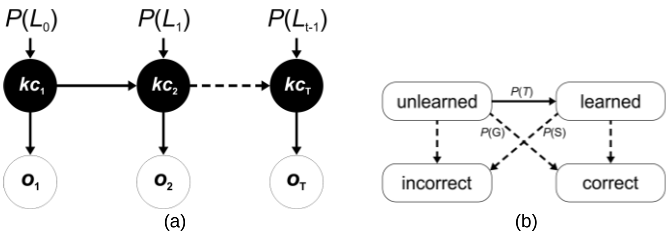

Figure 3. BKT model representation (a) BKT model framework and (b) relationship between bisection knowledge states and bisection performance

Additionally, equations (4) and (5) are used in the evaluation stage to calculate RMSE and MAE. According to the equation, c_i represents the ground truth score and p_i represents the predicted score [25]. The RMSE provides a summary of the model's prediction error, with a lower value indicating better performance [20]. This study also employs MAE for the purpose of regression. Similar to RMSE, lower MAE values denote better performance.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(c_i - p_i)^2} \tag{4}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|c_i - p_i| \tag{5}$$

III. RESULTS AND DISCUSSION

A. Student Performance Prediction

Error values resulting from performance predictions diminish in magnitude as the precision of the estimation of a student's knowledge state improves. This study uses the Moodle 2023 dataset, which was acquired and processed as described in the previous chapter. Furthermore, this analysis employs two renowned datasets, specifically Bridge to Algebra 2006–2007 from Cognitive Tutor and ASSISTments 2009–2010 from ASSISTments, for the purpose of comparison. Each of the three datasets includes student assessment records from time step 1 to t, allowing for the prediction of student performance at each time step. The results of student modeling using BKT on the three datasets are depicted in Figure 4. The graphic clearly demonstrates that the Moodle 2023 dataset produces the most favorable outcomes. The RMSE value obtained is 0.314, while the MAE value is 0.197. The complexity of the features in the Moodle 2023 dataset has been lowered compared to the Bridge to Algebra 2006–2007 and ASSISTments 2009–2010 datasets. The findings were also superior to those of the two earlier studies [9], [13] mentioned in the introduction.

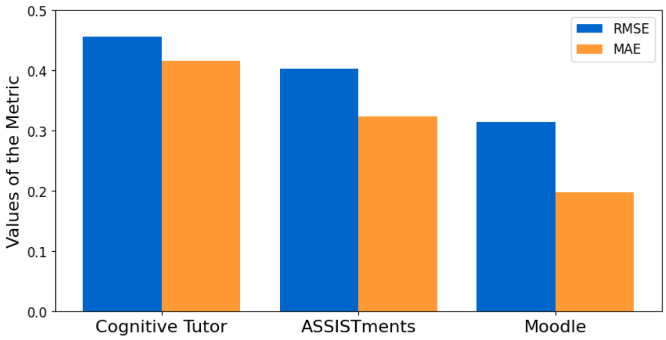

Figure 4. Results of all comparison datasets using BKT

B. Composition of Scores after Data Preprocessing

After passing through the data pre-processing stage, the composition score of each attempt was obtained, as shown in Figure 5. This image demonstrates the three problems that students should study: Tree Data Structure, Graph Data Structure, Searching & Sorting Algorithms. Among the three problems determined, it was seen that the majority of students obtained scores ranging from 0 to 84 on the Searching & Sorting Algorithms problem. As a result, we are conducting deeper investigation into the Searching & Sorting Algorithms problem. A total of 166 attempts were made, with 130 being the students' first attempt and 36 being their second attempt. When solely considering the quiz results in terms of scores, educators can only infer that a significant number of students fail to achieve the minimum score in their first and second attempts. However, it remains unclear which specific skills (KCs) pose challenges for students and how they progress on each question. Hence, the use of the BKT model can offer a potential solution to these issues by providing insights into students' performance and knowledge state. The subsequent chapter discusses and visualizes the applicability of the BKT model, particularly with regard to Searching & Sorting Algorithms problems.

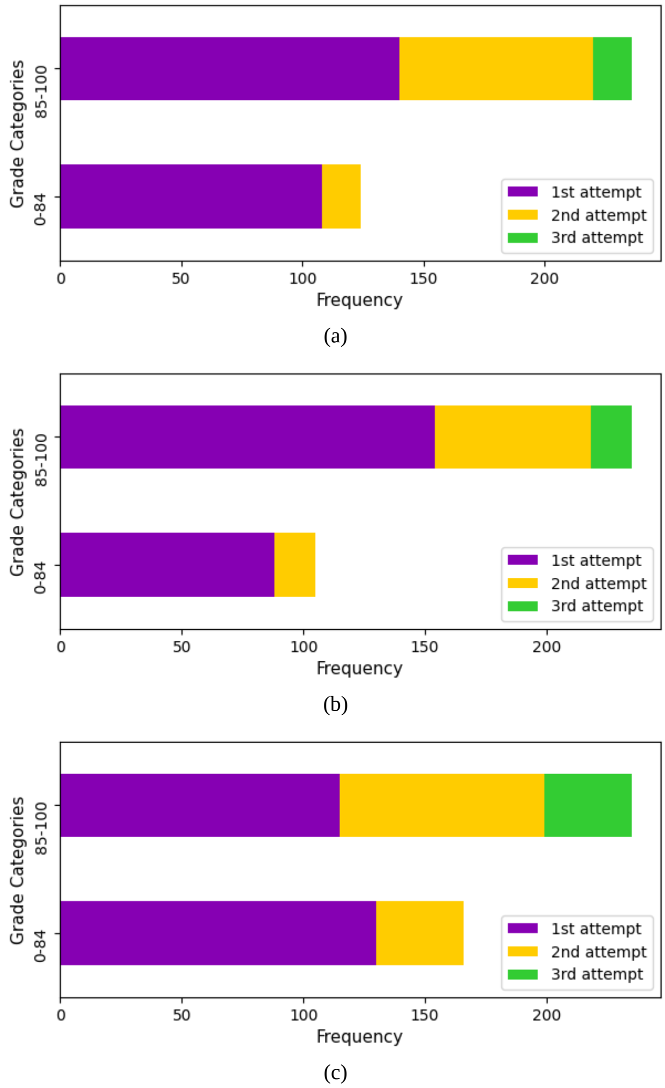

Figure 5. Composition score for three problems (a) Tree Data Structure, (b) Graph Data Structure, and (c) Searching & Sorting Algorithms

C. Visualization of Student Performance Predictions and Knowledge State

Figure 6 depicts predictive student performance movements and their knowledge state in the Searching & Sorting Algorithms problem with the knowledge component, specifically linear and binary search. This study includes 23 KCs from three problems; however, for the sake of the brevity of this article, we show and discuss only a set of knowledge components, i.e., linear and binary search. As with the discussion regarding predicted results from student performance, we chose three students who had maximum, middle, and minimum performance to be explained in this manuscript. On the left image, a gray line represents actual students' answers, and a blue line represents the outcome of the BKT prediction, which describes the predicted student performance. Blue lines on the right-hand image indicate the degree of the students' mastery of the KC. The image depicts three students (i.e., S_1 for the first student, S_2 for the second student, and S_3 for the third student) with varying levels of

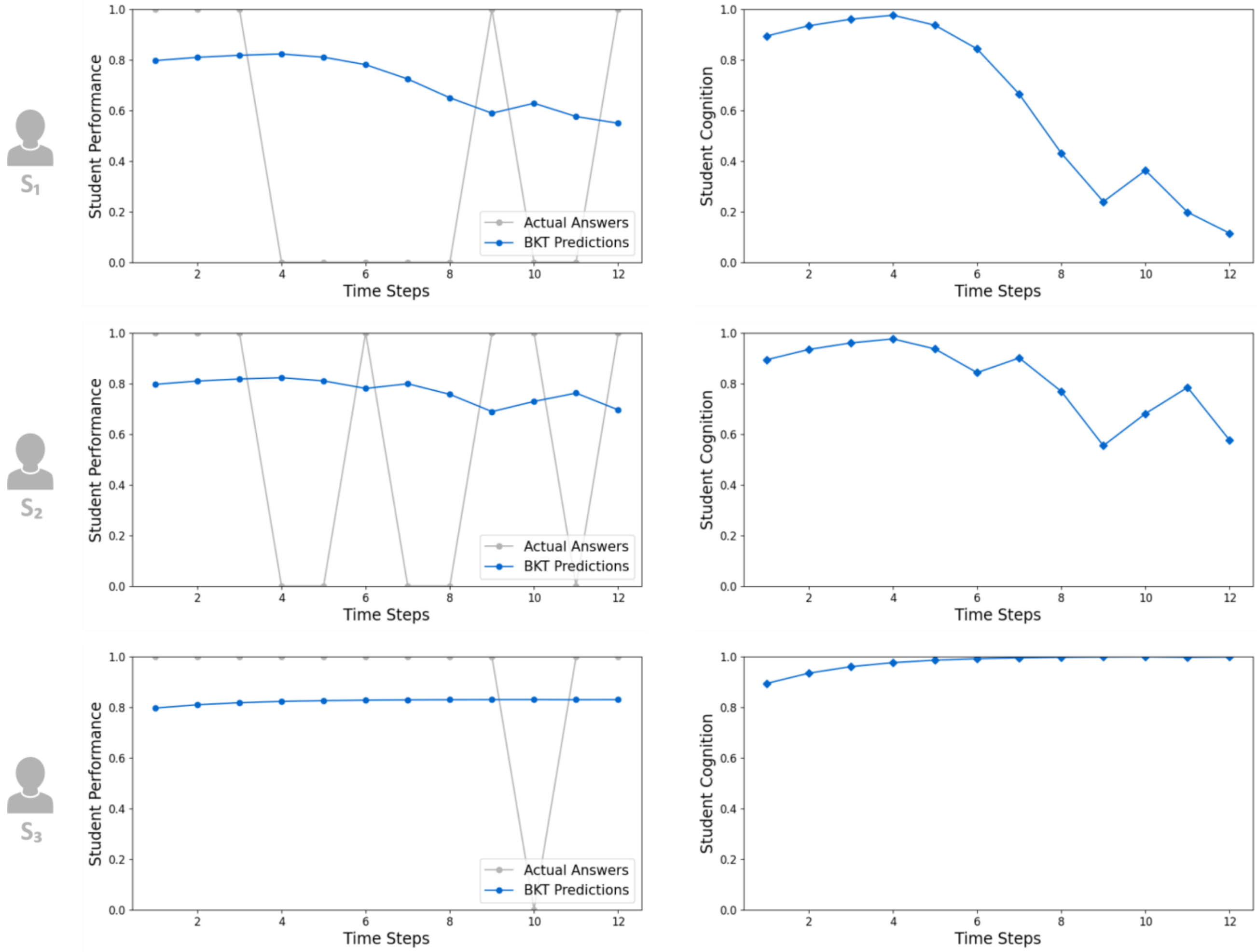

Figure 6. Performance predictions of three students and their knowledge state on the KC linear and binary search

mastery in the linear and binary search knowledge component. The blue lines will serve as the focal point of discussions pertaining to student performance predictions and knowledge states.

The level of mastery of the first student (S_1) was the lowest. Predictions of declining student performance serve to confirm this even further. The performance of S_1 decreased significantly (from 0.822 to 0.589) at time steps 4 to 9 due to incorrect answers, which contributed to a decline in student cognition (from 0.976 to 0.239). The student's performance continued to decline (time steps 11 and 12) until his or her cognition on KC's linear and binary search reached 0.114.

Meanwhile, the second student (S_2) has a moderate level of mastery. There are several up-and-down movements measured from 0 to 1, not only 0 or 1. This is an intriguing aspect of the performance prediction for S_2. Although S_2 has exhibited several performance decreases (time steps 5, 6, 8, 9, and 12), they do not occur sequentially as they do for S_1. Consequently, S_2 shows no significant cognitive decline. This KC's proficiency of S_2 is 0.576.

The third student (S_3) has a high level of mastery. The image shows that the performance is progressively rising. Despite the incorrect answer at the 10^{th} time step, there was no significant decline in the performance of S_3 (from 0.8299 to 0.8293). This resulted in a slight rise in the students' cognition, from 0.894 to 0.998.

The results and analysis that have been presented demonstrate that BKT can aid the process of searching for knowledge by producing predictions that allow for the estimation of changes in the knowledge state of the student. It can facilitate the implementation of blended learning. Due to the online nature of the assessment process, educators are unable to directly observe the constraints encountered by students during the execution of a series of exercises. Through these results, educators are able to estimate the specific level of students' knowledge at each KC and time step.

Apart from the advantages of the proposed knowledge tracing model, there are several things that still pose challenges. The most important thing is the quality of the dataset to ensure the effectiveness and reliability of the model. Because the dataset is obtained from recording student activities during the learning process, the setting of the learning environment is the key to success. This research only explores one of the features of Moodle, namely quizzes. There are many other features that can be utilized to better capture the development of student abilities over time. For example, different types of questions can provide various assessment data, like essays, videos, and audio. Further, the granularity of the knowledge component may be increased to allow for a more detailed understanding of the learner's progress. In addition, the dataset can be collected over a longer period of time to facilitate a more in-depth understanding of students' progress over time. By incorporating these improvements into a dataset for knowledge tracing, researchers and practitioners can develop more effective models for understanding and predicting students' knowledge states.

D. The Impact of BKT Implementation in Moodle

The optimization of Moodle's performance as a platform for blended learning environments can be achieved through the utilization of the BKT model. Educators can enhance their understanding of students' levels of knowledge by employing predictive techniques and afterward analyzing the resulting visualizations. This enables educators to ascertain the specific KCs that students have successfully acquired. The BKT model facilitates the process of knowledge tracing in a large class with many students. The model's outputs can be applied to learning analytics to monitor students' knowledge state. Learning analysis offers several advantages, one of which is the ability to personalize the materials, interventions, instructions, and exercises delivered to students based on their existing skills and abilities. Students' engagement, motivation, and self-direction will increase as a direct outcome of the implementation of personalized learning.

IV. Conclusion

The BKT model for knowledge tracing can be implemented in a blended learning environment. The results of RMSE 0.314 and MAE 0.197 support the statement. Predictions of student performance and estimates of student cognition can help educators understand student progress on each existing KC. In a blended learning environment, many students can be estimated at the same time. However, there are some areas that require further development. The knowledge-tracing process could be optimized by maximizing the settings on Moodle as a blended learning platform, allowing for more accurate predictions of student performance. In addition, the knowledge component's granularity can be increased for a more detailed understanding of learner progress, while a longer dataset collection time can provide in-depth insights.

References

[1] X. Li, Y. Yang, S. K. W. Chu, Z. Zainuddin, and Y. Zhang, "Applying blended synchronous teaching and learning for flexible learning in higher education: an action research study at a university in Hong Kong," *Asia Pacific J. Educ.*, vol. 42, no. 2, pp. 211–227, 2022, doi: 10.1080/02188791.2020.1766417.

[2] L. J. López-Reyes, A. L. Jiménez-Gutiérrez, and D. Costilla-López, "The Effects of Blended Learning on the Performance of Engineering Students in Mathematical Modeling," *Educ. Sci.*, vol. 12, no. 12, 2022, doi: 10.3390/educsci12120931.

[3] H. C. Wei, Y. H. Lin, and L. H. Chang, "The Effectiveness of a Blended Learning-Based Life Design Course: Implications of Instruction and Application of Technology," *SN Comput. Sci.*, vol. 4, no. 4, pp. 1–13, 2023, doi: 10.1007/s42979-023-01730-3.

[4] D. Evenhouse, Y. Lee, E. Berger, J. F. Rhoads, and J. DeBoer, "Engineering student experience and self-direction in implementations of blended learning: a cross-institutional analysis," *Int. J. STEM Educ.*, vol. 10, no. 1, pp. 1–22, 2023, doi: 10.1186/s40594-023-00406-x.

[5] D. Adinda and N. Mohib, "Teaching and instructional design approaches to enhance students' self-directed learning in blended learning environments," *Electron. J. e-Learning*, vol. 18, no. 2, pp. 162–174, 2020, doi: 10.34190/EJEL.20.18.2.005.

[6] C. C. Y. Yang and H. Ogata, "Personalized learning analytics intervention approach for enhancing student learning achievement and behavioral engagement in blended learning," *Educ. Inf. Technol.*, vol. 28, no. 3, pp. 2509–2528, 2023, doi: 10.1007/s10639-022-11291-2.

[7] E. Ameloot, T. Rotsaert, and T. Schellens, "The supporting role of learning analytics for a blended learning environment: Exploring students' perceptions and the impact on relatedness," *J. Comput. Assist. Learn.*, vol. 38, no. 1, pp. 90–102, 2022, doi: 10.1111/jcal.12593.

[8] J. Mandlazi, A. Jadhav, and R. Ajoodha, "Educational data mining: Using knowledge tracing as a tool for student success," *2021 3rd Int. Multidiscip. Inf. Technol. Eng. Conf. IMITEC 2021*, pp. 1–7, 2021, doi: 10.1109/IMITEC52926.2021.9714608.

[9] L. Meng, M. Zhang, W. Zhang, and Y. Chu, "CS-BKT: introducing item relationship to the Bayesian knowledge tracing model," *Interact. Learn. Environ.*, vol. 29, no. 8, pp. 1393–1403, 2019, doi: 10.1080/10494820.2019.1629600.

[10] H. Wan, Z. Zhong, L. Tang, and X. Gao, "Pedagogical Interventions in SPOCs: Learning Behavior Dashboards and Knowledge Tracing Support Exercise Recommendation," *IEEE Trans. Learn. Technol.*, vol. 16, no. 3, pp. 431–442, 2023, doi: 10.1109/TLT.2023.3242712.

[11] G. Abdelrahman, Q. Wang, and B. Nunes, "Knowledge Tracing: A Survey," *ACM Comput. Surv.*, vol. 55, no. 11, pp. 1–37, 2023, doi: 10.1145/3569576.

[12] Y. Mao, C. Lin, and M. Chi, "Deep Learning vs. Bayesian Knowledge Tracing: Student Models for Interventions," *J. Educ. data Min.*, vol. 10, no. 2, pp. 28–54, 2018.

[13] K. Zhang and Y. Yao, "A three learning states Bayesian knowledge tracing model," *Knowledge-Based Syst.*, vol. 148, pp. 189–201, 2018, doi: 10.1016/j.knosys.2018.03.001.

[14] Z. Jin, K. Ma, K. Liu, and K. Ji, "Exercises Recommendation in Adaptive Learning System," in *Proceedings of the 2nd International Conference on Big Data Technologies*, 2019, pp. 97–100.

[15] X. Ding and E. C. Larson, "Why deep knowledge tracing has less depth than anticipated," *EDM 2019 - Proc. 12th Int. Conf. Educ. Data Min.*, no. Edm, pp. 282–287, 2019.

[16] S. Pandey and G. Karypis, "A self-attentive model for knowledge tracing," *EDM 2019 - Proc. 12th Int. Conf. Educ. Data Min.*, pp. 384–389, 2019.

[17] X. Ding and E. C. Larson, "Incorporating uncertainties in student response modeling by loss function regularization," *Neurocomputing*, vol. 409, pp. 74–82, Oct. 2020, doi: 10.1016/j.neucom.2020.05.035.

[18] W. Gan, Y. Sun, X. Peng, and Y. Sun, "Modeling learner's dynamic knowledge construction procedure and cognitive item difficulty for knowledge tracing," *Appl. Intell.*, vol. 50, no. 11, pp. 3894–3912, 2020, doi: 10.1007/s10489-020-01756-7.

[19] M. Khajah, R. V Lindsey, and M. C. Mozer, "How Deep is Knowledge Tracing?," *arXiv preprint arXiv:1604.02416*, 2016.

[20] S. Minn, "AI-assisted knowledge assessment techniques for adaptive learning environments," *Comput. Educ. Artif. Intell.*, vol. 3, no. July 2021, p. 100050, 2022, doi: 10.1016/j.caeai.2022.100050.

[21] F. Liu, X. Hu, C. Bu, and K. Yu, "Fuzzy Bayesian Knowledge Tracing," *IEEE Trans. Fuzzy Syst.*, vol. 30, no. 7, pp. 2412–2425, 2022, doi: 10.1109/TFUZZ.2021.3083177.

[22] Q. Liu, S. Shen, Z. Huang, E. Chen, and Y. Zheng, "A Survey of Knowledge Tracing," *IEEE Trans. Knowl. Data Eng.*, pp. 1–20, 2021.

[23] A. Abyaa, M. Khalidi, and I. Samir, "Learner modelling : systematic review of the literature from the last 5 years," *Educ. Technol. Res. Dev.*, vol. 67, no. 5, pp. 1105–1143, 2019, doi: 10.1007/s11423-018-09644-1.

[24] R. Pelánek, "Bayesian knowledge tracing, logistic models, and beyond : an overview of learner modeling techniques," *User Model. User-adapt. Interact.*, vol. 27, no. 3, pp. 313–350, 2017, doi: 10.1007/s11257-017-9193-2.

[25] J. Zu, Q. Li, J. Lu, P. LV, and G. Yu, "Leveraging Cognitive Diagnosis to Improve Peer Assessment in MOOCs," *IEEE Access*, vol. 9, pp. 50466–50484, 2021, doi: 10.1109/ACCESS.2021.3069055.