Enhancement of Prediction Model for Students' Performance in Intelligent Tutoring System

Hana Fakhira Almarzuki,[1] **Khyrina Airin Fariza Abu Samah,**[1] **Siti Khatijah Nor Abdul Rahim,**[2] **Shafaf Ibrahim,**[2] **and Lala Septem Riza**[3]

[1]College of Computing, Informatics and Mathematics, Universiti Teknologi MARA Cawangan Melaka Kampus Jasin, Melaka, Malaysia

[2]College of Computing, Informatics and Mathematics, Universiti Teknologi MARA Shah Alam, Selangor, Malaysia

[3]Department of Computer Science Education, Universitas Pendidikan Indonesia, Bandung, Indonesia

Abstract: Adapting artificial intelligence to intelligent tutoring system (ITS) has made teaching and learning more effective. Prediction of students' performance has gained more interest among researchers to know whether the students master their learning before moving to another topic. For the research scope, we have analyzed numerous Bayesian Knowledge Tracing (BKT) variations in methodology and found that the most precise way to forecast students' success is through Individualized Bayesian Knowledge Tracing (iBKT). Although iBKT makes a good prediction, iBKT does not consider other knowledge-related elements, such as learning and guess rate and only uses students' prior knowledge as the parameters. Due to issues concerning uncertainties in students' interactions, this study proposes to enhance the prediction function of the iBKT using a feature relating to students' confidence levels. Thus, this new confidence parameter is defined as P(C), assumed to improve prediction accuracy when forecasting student achievement. The prediction accuracy is tested using the attributes of the ASSISTment and Knowledge Discovery and Data Mining (KDD) datasets as input. In addition, root mean square error (RMSE) is applied to calculate the performance accuracy of iBKT and enhanced iBKT with the confidence parameter. As a result, the RMSE performance accuracy of iBKT with the confidence parameter shows a low RMSE score for both datasets. The ASSISTment dataset provides a higher prediction when applying the confidence parameter, 0.21190. Therefore, it is concluded that enhancing the confidence parameter is an effective method with accuracy improvement for predicting students' success in ITS.

Keywords: enhancement prediction; intelligent tutoring system; student model; students' performance

I. INTRODUCTION

There are two types of learning, which are traditional learning and online learning. Traditional learning is a learning method in which the learning content is taught by an instructor in a classroom at a prescribed time [1]. The rapid evolution of technology has paved the way for creating technologies used for learning, collectively referred to as e-learning. E-learning brings many main concepts, such as web-based learning practices. Online learning has become one of the most important e-learning platforms worldwide. With the spread of the COVID-19 pandemic, ensuring the continuity of student academic development has become even more essential. An intelligent tutoring system (ITS) is one type of e-learning system. Crowston *et al.* [2] claimed that ITS could increase the effectiveness of an individual tutoring system through digital resources. Moreover, the artificial intelligence (AI) aspect enables reactions as human discourse through inquiries of related elements and education [3]. It helps learners or students gain a good understanding effectively during their learning.

The key features of any tutoring system are the capability of gathering and modeling student data based on the response, motivation or emotion, metacognition, and instruction adjustment according to individual requirements [4]. The student's performance can be predicted with the help of AI in ITS. It is crucial to provide individualized tutoring based on the student's current mastery level or knowledge state, which can predict how well the students will perform in their courses.

ITS has four modules: expert module, tutoring module, student module, and interface module [5]. This study focuses on the student module where the prediction of the students' performance happened. The main objective of this research is to focus on predicting students' performance in the ITS. It is a computer-based educational system that applies AI concepts intended to improve students' performance in the learning process. The important task in a computer-based educational system is modeling and predicting students' performance [6]. The prediction in the student model will measure the student's knowledge level, which can improve student performance. Therefore, it is crucial to develop an accurate student model in order to determine the learning material and the level of student knowledge [7].

Though many suitable methods exist for making predictions in ITS, the most well-known state-of-the-art methods are Bayesian Knowledge Tracing (BKT) methods [8]. BKT is a common model

---
Corresponding author: Khyrina Airin Fariza Abu Samah (e-mail: khyrina783@uitm.edu.my).

used to predict students' knowledge individually and the probability that they can answer the question correctly or not based on their skill. BKT is the most used method for sequencing questions for students to enhance performance and engagement [9]. This approach follows the mastery learning paradigm, providing students with the questions expected to advance their learning, for example, the question that involves critical thinking. It proves that this method follows the mastery learning paradigm as it gives questions from a common question to a high-level one, and the students will still be able to solve it.

Despite the excellent performance demonstrated by the BKT methods, there are still problems in accurately measuring and predicting students' performance [10]. It is due to the limitation that exists in the method used itself. One of the BKT's limitations is that it only focuses on skills' parameters, which refer to whether the students have mastered the learning or not [11]. Thus, extending BKT's method by adding a confidence parameter could help enhance student performance prediction in ITS [12].

This study analyses several methods to predict the students' performance. After analyzing it, the best method was chosen and further enhanced to predict the performance of a set of students. For studying the purpose prediction accuracy, the result generated by the enhanced method will be compared with the result generated by the original method before the enhancement.

Therefore, as a result, we propose to enhance the prediction function of the iBKT. A new parameter, confidence, is added related to the student's confidence levels. We assumed that the confidence parameter can improve the prediction accuracy in forecasting student performance.

The remainder of this paper is organized as follows. Section II reviews the ITS and its Methods in Measuring and Predicting Student Performance in the literature. It also presents the insights gained from the extensive reviews. Section III elaborates on the proposed methodology, which is the enhancement and implementation of iBKT using two types of datasets. Section IV discusses the results obtained through the experiments done by employing the proposed method. Finally, this study's conclusion is presented in Section V.

II. LITERATURE REVIEW

A. INTELLIGENT TUTORING SYSTEM

ITSs are sophisticated, automated software systems that apply AI concepts and methods to teaching and learning problems and needs [13]. ITS is an automated system that can provide a personalized tutor for learning [14]. It has four modules, as in Fig. 1, that consist of the expert module, tutoring module, student module, and interface module [5]. The expert model is used to bring together all the information relevant to the field of expertise required for the teaching process. This module must generate questions, explanations, answers, hints, and comments and provide standards to check on the student's performance.

The student's module must be represented by students' answers, behaviors, and actions, which means the student's knowledge and skills [15]. Furthermore, ITS does not work without knowing the student's knowledge, so awareness, behavior, and all other aspects that could affect the student's performance should be included in this module. The Tutor or Pedagogical Model puts in place pedagogical techniques to teach the knowledge of a given area. Any educational approach must be based on the psychological
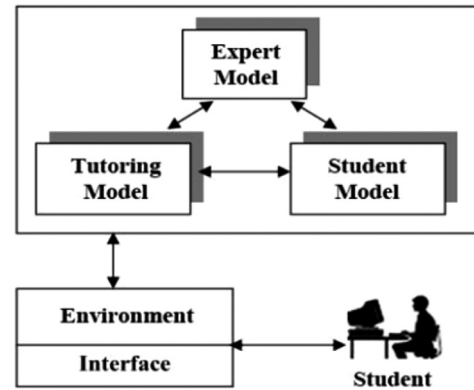

Fig. 1. Module in the intelligent tutoring system.

and pedagogical concepts of type. The learner's needs must be defined and expressed in the teaching approach.

The interface model helps the students to communicate effectively. It involves thinking about user reasoning, handling communication, sketching graphics to demonstrate argument, displaying or detecting emotions, and describing how conclusions were reached [16]. All of these modules are important in developing the ITS.

ITS offers many benefits, one of which is the power to adjust to students' learning styles so that students can learn at their own pace and according to their preferences [17]. The main role of the ITS is to assess the mastery learning of the student. Motivated by the issues that have attracted much attention from researchers working on ITS concerning prediction accuracy and students' uncertainties, this study focuses on the student module. The performance prediction is based on the individual student. It is to measure the student's mastery level of the underlying knowledge. It also determines the student's knowledge about a specific field and how accurately and correctly a student answers the task based on their responses [18].

B. ITS METHODS IN MEASURING AND PREDICTING STUDENT PERFORMANCE

Knowledge Tracing (KT) has gained a reputation in this field of study for providing effective guidance in the learning process. KT aims to determine learned skills and use this knowledge to customize learning opportunities in applications such as tutoring programs, on the web, and even in Massive Open Online Courses [19]. By doing this, mastering a skill can be achieved. KT aims to track students' knowledge state based on observed results from their previous practices [19]. It also can be known as student modeling. The most popular method in KT is Bayesian Knowledge Tracing [8,20].

BKT is a method to estimate mastery learning. One of the advantages of the BKT model is that it consists of both theoretical roots and efficient real-world implementation [21]. The function of the BKT model is used to track the probability of whether the student has mastered the learning and the student can answer the questions correctly [22]. The prediction of the student's skill is based on the given set of assessments to ensure the student has mastered the learning and is able to move to the next level of learning [23].

Since BKT is the most popular method and is good at predicting mastery learning, many researchers have used it to predict student performance. Even though BKT shows a good performance, it limits on focusing on skill parameters; thus, the researcher has introduced the extended version of BKT, which is Individualized Bayesian Knowledge Tracing (iBKT) [24], Student Skill Model (SSM) [25], and Personalized, Clustered, Bayesian Knowledge Tracing (PC-BKT) [21]. Other than BKT, Performance Factor Analysis (PFA) is the other method of KT [26]. It is a new alternative to predicting student performance based on reconfiguring Learning Factor Analysis [27].

Therefore, developing an accurate method to predict students' performance in ITS is important. The extensive review aims to identify which method can provide the most accurate prediction of students' performance. Table I summarizes the in-depth review of the original BKT method used in approximate mastery learning. It might result in less prediction accuracy since it concentrates more on skill parameters [22]. Thus, researchers in the literature have developed a few improvements to BKT approaches to improve student performance prediction in the ITS. Several improvised BKT techniques have been researched and compared, including the iBKT, SSM, PC-BKT, and PFA.

The original BKT provides less prediction accuracy compared to the enhanced iBKT. Thus, adding confidence parameters to improve and get better prediction accuracy is worthwhile. The performance of iBKT, SSM, and PC-BKT regarding prediction accuracy is practically on par. These three techniques were successful in outperforming the original BKT. The PFA reportedly performed quite well, although some tweaking parameters are required to perform an accurate prediction. Additionally, it was claimed that PFA and BKT are equivalent in terms of prediction accuracy when it comes to multi-skill questions.

According to the comparative research, PFA is less effective than the others. In order to choose the optimal approach to use in our research, PFA excludes other methods and concentrates more on iBKT, SSM, and PC-BKT. iBKT does not consider other knowledge-related factors like learning and guess rate and only employs student-prior knowledge as the parameters. Four student and four skill parameters are employed simultaneously in the SSM, which uses more parameters to make predictions.

In the case of the PC-BKT, its prediction function includes the "forgetting" component. PC-BKT is not used in our research since we want to concentrate on predicting students' performance on the mastery level of a subject usually finished in just one or two weeks. Additionally, researchers asserted that if a student does not use the skill within 30 days, there is a greater chance that they will "forget" it. We anticipate that the student guessing and learning rates can be incorporated into the iBKT to improve the prediction function. We know that the iBKT generates a good prediction despite not considering these factors. It will be fascinating to see if the newly proposed iBKT can outperform the SSM, which already uses all the components as parameters in its prediction function.

The confidence level analysis is crucial since it reveals whether students guess during assessments and can identify problems with student uncertainty in learning. According to a recent study [33], confidence level should be considered when evaluating a student's performance to reduce uncertainty in the interaction of the students and improve the student model in the ITS. It is proven that a high percentage of the students in the dataset were predicted correctly. Thus, it is hypothesized that a proposed method with the addition of confidence as an adding parameter could enhance the accuracy prediction of student's performance in ITS.

We employed iBKT with confidence parameters in our next research stage. The aim is to determine the uncertainty and enhance the student's performance prediction. The performance of the prediction accuracy was compared between the initial iBKT and the enhanced iBKT.

III. MATERIALS AND METHODS

This section presents the prediction step based on the method analyzed during the literature review, which is the iBKT method. Figure 2 shows the study flow identified the steps involved in predicting the student's performance. It consists of four steps in general with specific methods involved.

A. DATA COLLECTION

We collect the data using two datasets: ASSISTment and Knowledge Discovery [34] and Data Mining (KDD) [35]. It ensures that the results produced using the proposed method are consistent, thus ensuring its reliability and universality. The ASSISTment is a free online web-based tutoring framework for 4th- to 10th-grade mathematics and collected data throughout the 2009–2010 school year. This dataset, which was generated in 2015, consists of 708632 problems from 100 skill-builder sets. The data focus on skill builders is also known as data on mastery learning. This dataset consists of problem sets for skill builders (mastery learning), where a learner is regarded to have mastered a skill once they reach a particular level.

Meanwhile, KDD Cup 2010 is data on the Information Discovery and Data Mining Challenge about student performance evaluation. It is an educational data mining competition in which participants are charged with predicting student algebraic problem performance given information regarding past performance. Specifically, participants were provided with summaries of the logs of student interaction with ITSs.

Datasets will be retrieved in this step and organized into a standardized type. A standardized dataset is essential to ensure a clear understanding of how the data are interpreted. Figure 3 shows the sample of the ASSISTment dataset, and Fig. 4 shows the sample of the KDD dataset. The collected dataset shows more than 1,000 students who worked on 708,632 problems.

For this research, since the original dataset is huge, first, the dataset is standardized into a simple form of a dataset, purposely to make it easier and more systematic to enhance the efficiency of predicting student performance in ITS. Thus, the first 20 students will be selected from the standardized dataset.

Based on Fig. 5, the data that will be standardized are in column K, which represent the answer type. In this research, the answer type will focus on multiple-choice questions. Therefore, the answer type will be standardized to "choose_1," representing the multiple-choice question.

B. DATA PREPROCESSING

The following process is preprocessing data. This step is essential to select the important attributes and focus areas to predict the students' performance. This research will focus on multiple-choice questions from the course and select 20 students from each dataset. The standardized form of the dataset in this step is preprocessed in order to select the important attributes.

Table II shows the sample of the preprocessed ASSISTment dataset, which consists of seven columns with the important label

Table I. ITS prediction using different methods

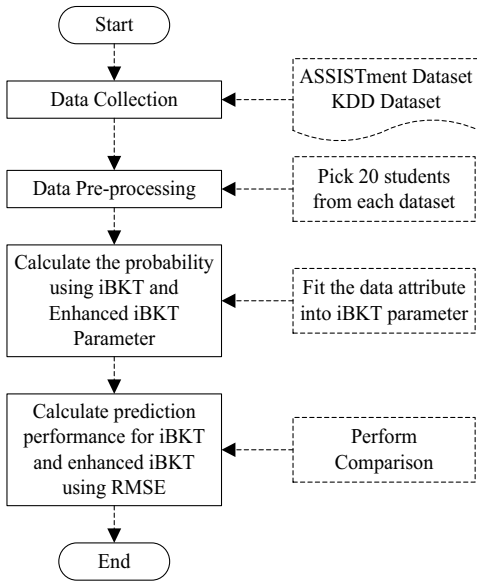| Characteristics	Methods				
	BKT	iBKT	SSM	PC-BKT	PFA
Parameter	Four parameters: • P(L0): Leaner (the probability of a student having the skill even before trying the lesson) • P(T): Transit (the probability be able to learn it after completing the lesson by a student who did not know the skill) • P(G): Guess (the probability student answering a question correctly in the lesson without knowing the skill) • P(S): Slip (the probability of a student having the skill but yet reacting incorrectly)	• In the iBKT models, only the P(L0) and P(T) were individualized. • The four parameters in BKT which, also known as 4N divided into two components: a component per skill and a component per student. • The number of parameters was 4N + 2M, where M is the number of students.	• Eight parameters simultaneously: Four student parameters and four skill parameters. • The number of parameters is 4N + 4M, where N refers to student and M refers to skill.	• This model assumes that over time, the amount of material learned that is remembered decays exponentially. $N(t) = N_{(t-30)}\, e^{-\lambda(t/30)}$ where N is the degree of information, lambda is an exponential constant of decay set to 0.1, and the interval of time is t. Assumption that if a student does not try the ability within 30 days, the risk of forgetting will increase.	• PFA is a variant decomposition of learning and relies on the reconfiguration of learning factor analysis (LFA). • By dropping the student variable and replacing the skill variable with the question identity, it can reconfigure LFA on its independent variable, for example, one parameter per question. Its model estimates a parameter that represents the difficulty each of the items and two parameters reflecting the effects of previous failures achieve for that each of skill.
Advantages	BKT is a method to approximate mastery learning. The prediction of the student skill is based on the given set of assessment to ensure the student has mastery the learning and be able to move to the next level of learning [29].	The individualization of BKT is done by insertion of a student node. The conditional probability table for the initial knowledge node is then dependent on the value of the student node [24].	This model can be used to keep track of all the four student parameters and four skill parameters simultaneously in a single process [25].	PC-BKT assumes that a skill once learned is never forgotten [21].	PFA can directly handle with multi-skill performances [28].
Disadvantages	The BKT method only focuses on the skill parameter, which refers to the student having mastered or unmastered the learning [29].	The individualized BKT method is only limited to students' prior knowledge. However, it does not extend to other aspects of student knowledge such as guess and learning rates [25].	The Student Skill Model can only be used when there is a large number of skills and students [30].	Challenging in gaining data using a sparse matrix [31].	PFA could generate a negative learning rate if it is overfitting [32].
Dataset	Cognitive Tutor dataset KDD	Cognitive Tutor KDD	ASSISTments	Cognitive Tutor KDD	ASSISTment
Prediction accuracy	Compared to PC_BKT and other enhancements of the BKT method, the BKT method is less accurate compared to other methods.	Proven that the implementation of the iBKT model can tangibly improve the accuracy by adding the student-specific probability of learning in predicting the success of student work in an ITS.	The average RMSE shows that the Student Skill Model is more accurate than KT, which is 0.4199 compared to 0.4331.	PC-BKT model that incorporates forgetting as a time decay function proves that the decay function increases the performance accuracy by 95.9%.	PFA is at par with KT's prediction, which deals with multi-skill questions. It can be improved with minor tweaking parameters (non-negative restricting learning rates) and KT parameter model fitting.

Fig. 2. Flowchart of processes to predict the students' performance.

- P(G): Guess (the probability of the student answering a question correctly in the lesson without knowing the skill).

- P(S): Slip (the probability of a student having the skill but yet reacting incorrectly).

Since there is a problem with this method in predicting student performance in ITS, the enhancement of the BKT method, which is iBKT, is used to enhance the prediction accuracy of the student performance in ITS.

We solve the uncertainty issue concerning ITS in student interaction by introducing the students' confidence criteria as one of the parameters defined as P(C). As for the mentioned dataset, the probability of a student answering a question correctly within the given time limit will be referred to.

The time recorded, "Ms first response," or the first response millisecond in the ASSSITment dataset and "Correct Step Duration (sec)" in the KDD dataset used for the selected students in answering a question. The number of attempts (in a time period) to answer the question can be used to measure the confidence level of the students. The P(C) parameter will be implemented in the iBKT to predict student performance.

to be analyzed. The data focus on the multiple-choice question and the same Template_id, which is 50,299. It indicates a similar question for the assignment assigned to the assignment students. Table III shows the sample of the preprocessed KDD dataset, which consists of nine columns with an important label to be analyzed.

Datasets retrieved and preprocessed are analyzed to identify which data will be set as parameters according to the BKT method. The Bayesian network learns the model's parameters and predicts the students' performance. In general, there are four parameters in BKT, which are P(L0), P(T), P(G), and P(S).

- P(L0): Leaner (the probability of a student having the skill even before trying the lesson).

- P(T): Transit (the probability of being able to learn it after completing the lesson by a student who did not know the skill).

C. CALCULATE THE PROBABILITY USING iBKT AND ENHANCED iBKT PARAMETER

Analyzing the dataset to fit the data attribute as an input parameter to the prediction function is crucial. Our study's prediction function, based on the ASSISTment dataset and four BKT parameters, is represented by Equations (1)–(4). Equation (1) can be interpreted as the proportion of students correctly answering on their first try without assistance, demonstrating that they understood the subject. The number of students who correctly answer the same question a second time is represented by Equation (2). The student does not know the lesson when they answer the question poorly the first time and ask for a hint, but they do after receiving the hint.

The number of students who attempt the question multiple times is shown in Equation (3). If the student repeatedly alters his response without understanding the lesson, the answer may be

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	order_id	assignment_id	user_id	assistment_id	problem_id	original	correct	attempt_count	ms_first_response	tutor_mode	answer_type	sequence_id	student_class_id
2	33022537	277618	64525	33139	51424	1	1	1	32454	tutor	algebra	5948	132
3	33022709	277618	64525	33150	51435	1	1	1	4922	tutor	algebra	5948	132
4	35450204	220674	70363	33159	51444	1	0	2	25390	tutor	algebra	5948	118
5	35450295	220674	70363	33110	51395	1	1	1	4859	tutor	algebra	5948	118
6	35450311	220674	70363	33196	51481	1	0	14	19813	tutor	algebra	5948	118
7	35450555	220674	70363	33172	51457	1	1	1	16031	tutor	algebra	5948	118
8	35450573	220674	70363	33174	51459	1	1	1	15047	tutor	algebra	5948	118
9	35480603	220674	70363	33123	51408	1	1	1	10732	tutor	algebra	5948	118
10	33140811	220674	70677	33168	51453	1	1	1	23241	tutor	algebra	5948	118
11	33140919	220674	70677	33112	51397	1	1	1	11512	tutor	algebra	5948	118
12	33275447	277618	70695	33149	51434	1	1	1	14891	tutor	algebra	5948	132
13	33275494	277618	70695	33188	51473	1	1	1	32844	tutor	algebra	5948	132
14	35222057	277618	70699	33145	51430	1	1	1	38860	tutor	algebra	5948	132
15	35222063	277618	70699	33151	51436	1	1	1	6249	tutor	algebra	5948	132
16	32871676	277618	70707	33136	51421	1	1	1	5710	tutor	algebra	5948	132
17	32871682	277618	70707	33111	51396	1	1	1	4741	tutor	algebra	5948	132
18	33140490	277618	70727	33123	51408	1	1	1	13816	tutor	algebra	5948	132
19	33140547	277618	70727	33117	51402	1	1	1	7441	tutor	algebra	5948	132
20	33392795	277618	70729	33169	51454	1	0	2	12898	tutor	algebra	5948	132
21	33392817	277618	70729	33175	51460	1	1	1	10955	tutor	algebra	5948	132
22	33392838	277618	70729	33140	51425	1	0	2	27857	tutor	algebra	5948	132

Fig. 3. Sample ASSISTment dataset.

Row	Anon Stude	Problem Hierarchy	Problem N	Proble	Step Name	Step Star	First Trai	Correct	Step End T	Step D	Correc	Error	Correc	Incor	Hints
9098	46z59pQP5	Unit CTA1_10, Secti	DISTFB06_S	1	0.3(X-100000)	57:07.0	57:16.0	57:16.0	57:16.0	9	9		1	0	0
9887	46z59pQP5	Unit ES_07, Section	LIT78A	6	di-expr	24:43.0	24:46.0	24:46.0	26:00.0	6	6		1	0	0
9932	46z59pQP5	Unit CTA1_12, Secti	LDEMO_SYS	2	R4C2	45:43.0	46:44.0		46:44.0	61		61	0	1	0
10039	52vEY7f17k	Unit CTA1_13, Secti	PROP01	1	2/300*X = 10	42:05.0	42:08.0	42:08.0	42:08.0	3	3		1	0	0
10482	52vEY7f17k	Unit CTA1_06, Secti	FOR05	1	xScale	39:38.0	39:41.0	39:41.0	39:44.0	6	6		1	0	0
11024	52vEY7f17k	Unit ES_03, Section	EG52	1	2.9 = -2x+3.7+	20:24.0	20:27.0	20:27.0	21:11.0	11	11		1	0	0
11476	52vEY7f17k	Unit CTA1_08, Secti	REAL10	1	yScale	28:36.0	28:38.0	28:38.0	28:38.0	2	2		1	0	0
11948	52vEY7f17k	Unit ES_04, Section	EG58	7	xR2	23:25.0	23:31.0	23:31.0	23:31.0	6	6		1	0	0
12532	52vEY7f17k	Unit CTA1_10, Secti	DIST05_SP	1	R1C2	26:00.0	26:05.0	26:19.0	26:19.0	13		13	0	1	0
12955	52vEY7f17k	Unit ES_07, Section	LIT69A	4	z*u-n*j+z*g =	50:30.0	51:04.0		51:04.0	34		34	0	0	1
13056	5ey2xUDZT	Unit CTA1_01, Secti	LDEMO_WH	1	R2C2	48:20.0	49:21.0	49:21.0	49:21.0	61	61		1	0	0
13516	5ey2xUDZT	Unit CTA1_14, Secti	2PTFB18	1	R7C2	31:14.0	31:20.0	31:20.0	31:20.0	6	6		1	0	0
13772	5ey2xUDZT	Unit LINEAR-INEQU	YGE4X-11	2	xMax	19:47.0	19:51.0	19:51.0	19:51.0	4	4		1	0	0
14059	5ey2xUDZT	Unit CTA1_15, Secti	SYS06	1	R5C3	39:39.0	39:40.0		39:46.0	0.7		0.7	0	2	0
14366	5ey2xUDZT	Unit QUAD-ADD-AR	NEWSQUAI	1	Point 1.1	31:13.0	31:27.0		32:09.0	56		56	0	0	29
14586	5ey2xUDZT	Unit QUADRATICS-F	EG-FACTOR	1	c-factor-1	26:54.0	28:02.0	28:02.0	28:02.0	68	68		1	0	0

Fig. 4. Sample KDD dataset.

A	B	C	D	E	F	G	H	I	J	K
32815144	277628	70695	51890	91214	1	1	1	34375	tutor	choose_1
32815179	277628	70695	51870	91194	1	1	1	19938	tutor	choose_1
32815190	277628	70695	52241	91570	1	1	1	7563	tutor	choose_1
32416542	277628	70699	51866	91190	1	0	3	79592	tutor	choose_1
34957398	277628	70727	51876	91200	1	1	1	67982	tutor	choose_1
35162291	277628	70729	51898	91222	1	0	1	39979	tutor	choose_1
35162306	277628	70729	52226	91555	1	0	2	21602	tutor	choose_1
32265959	277463	70872	51898	91222	1	0	4	200328	tutor	choose_1
32266649	277463	70872	51868	91192	1	1	1	33735	tutor	choose_1
32266989	277463	70872	52242	91571	1	1	1	15938	tutor	choose_1

Fig. 5. Sample of standardized ASSISTment dataset.

Table II. Sample of preprocessed ASSISTment dataset

Assignment Id	User Id	Correct	Attempt_count	Hint_count	Ms_first_response	Template_id
277628	70695	1	1	0	34375	50299
277628	70729	0	2	3	21602	50299
263371	77962	0	4	0	64048	50299
270284	78073	1	1	0	117427	50299
263771	78804	1	1	0	137349	50299

correct or inaccurate, showing that the learner is simply guessing the question. The percentage of students correctly responding to the first question on their first try is shown in Equation (4). The students gave the same response wrong the second time, indicating they might have made a careless error.

In order to more accurately anticipate the students' performance, $P(C)$ is added to the original function to improve the iBKT approach in our study. According to Equation (5), it can be described as the proportion of students who correctly answer the question the first time on their first try and do so quickly, showing that they are very confident. Next, Equations (6)–(8) are the predicting function based on the KDD dataset. The data assigned is the same term in Equations (1) and (2). Equation (8) is the number of students who answered the question incorrectly.

However, there is a slight chance the student will answer correctly by making a guess. In the KDD dataset, the student cannot afford to make a silly mistake as they can view the question or problem several times before answering the question, so $P(S)$ is not considered in the KDD dataset. Equation (9) is for enhancement of the iBKT method using the KDD dataset regarding confidence level. It is determined based on how many students answer the question in less step duration. However, if the student answers it correctly, we considered the student answers the question confidently.

Only two distinct priors of student information, which are high and low, can be used in the iBKT approach. It can make use of a

Table III.   Sample of preprocessed KDD dataset

Row	Step start time	First transaction time	Correct transaction time	Correct step duration (sec)	Correct first attempt	Incorrects	Hints	Corrects
9098	57:07.0	57:16.0	57:16.0	9	1	0	0	1
9887	24:43.0	24:46.0	24:46.0	6	1	0	0	2
10039	42:05.0	42:08.0	42:08.0	3	1	0	0	1
10482	39:38.0	39:41.0	39:41.0	6	1	0	0	2
11024	20:24.0	20:27.0	20:27.0	11	1	0	0	2

student's initial response. Students who respond incorrectly to their first question are placed in a low category. In addition, if they gave the correct answer, they thought they fell into a high category. Setting them as ad hoc is one potential strategy for forecasting it. The high prior ad hoc, conditional on the initial response, is justified by an intuition of about 1 minus the probability of a guess [36]. Low prior should match the probability slip in terms of value.

D. CALCULATE THE PREDICTION PERFORMANCE OF iBKT And Enhanced iBKT USING RMSE

Root mean square error (RMSE) is used to calculate the performance accuracy of iBKT and enhance iBKT by adding P(C). In particular, skill model evaluation frequently uses the RMSE statistic [37]. Any anticipated and actual values (observed values) are compared using RMSE. The iBKT method is on the RMSE presented in Equation (10) to measure iBKT's performance, and Equation (11) is used to measure the iBKT with the addition of P(C).

Following that, we compare the two RMSE values. A lower RMSE score is considered to have a higher degree of prediction accuracy [37]. The discrepancy between the predicted and actual values increases with increasing RMSE, indicating that the model is less accurate:

$$P(L0) = \frac{No\ of\ Students(correct, oneattempt, no\ hint)}{Total\ No\ of\ Students} \quad (1)$$

$$P(T) = \frac{No\ of\ Students(correct, hint)}{Total\ No\ of\ Students} \quad (2)$$

$$P(G) = \frac{No\ of\ Students(correct, manyattempt, nohint)}{Total\ No\ of\ Students} \quad (3)$$

$$P(S) = \frac{No\ of\ Students(incorrect, oneattempt)}{Total\ No\ of\ Students} \quad (4)$$

$$P(C) = \frac{No\ of\ Students(correct, oneattempt, shortmsfirstresponse)}{Total\ No\ of\ Students} \quad (5)$$

$$P(L0) = \frac{No\ of\ Students(correct)}{Total\ No\ of\ Students} \quad (6)$$

$$P(T) = \frac{No\ of\ Students(correct, hint)}{Total\ No\ of\ Students} \quad (7)$$

$$P(G) = \frac{No\ of\ Students(incorrect, correct\ step\ duration)}{Total\ No\ of\ Students} \quad (8)$$

$$P(C) = \frac{No\ of\ Students(Step\ Start, First\ transaction\ time, correct\ transaction, correctstepduration)}{Total\ No\ of\ Students} \quad (9)$$

$$Performance\ of\ iBKT = \sqrt{\frac{P(L0) + P(T) + (1 - P(G)) + P(S) - (actual\ value)}{Total\ No\ of\ Students}} \quad (10)$$

$$Performance\ of\ iBKT\ with\ P(C) = \sqrt{\frac{P(L0) + P(T) + (1 - P(G)) + P(C) + P(S) - (actual\ value)}{Total\ No\ of\ Students}} \quad (11)$$

IV. RESULT AND DISCUSSION

Analyzing the data from the dataset is important as it will be implemented in the prediction function to predict the student's performance. Two bar graphs were plotted to show the difference.

For clarification, the standard BKT approach, in this case, is the iBKT without P(C). In contrast, the improved iBKT method is the iBKT with P(C). Table IV shows the dataset result for each parameter from the equation in the methodology phase. As mentioned, the KDD dataset for P(S) is not considered. Thus, we mark

Table IV.   Dataset result for each parameter

Parameter	KDD dataset Value	ASSISTment dataset Value
P(L0)	0.6670	0.4670
P(T)	0.0000	0.1330
P(G)	0.0670	0.4670
P(S)	NA	0.0670
P(C)	0.2000	0.1330
iBKT	1.6000	1.2000
iBKT with P(C)	1.4000	1.0670
RMSE without P(C)	0.2826	0.2319
RMSE with P(C)	0.2590	0.2119

Fig. 7. RMSE value for dataset ASSISTment iBKT with P(C) and KDD iBKT with P(C).

iBKT without P(C) using dataset ASSISTment with 0.25060 and dataset KDD with 0.25062. Figure 7 depicts the RMSE value 0.21190 for dataset ASSISTment iBKT with P(C), and the RMSE value for dataset KDD iBKT with P(C) indicates 0.23519.

We analyze the result by focusing on a few angles, as in Table V. First, we compare the results based on a specific individual dataset, ASSISTment and KDD. RMSE result for dataset ASSISTment iBKT without P(C) shows 0.25060, which is higher compared to ASSISTment iBKT with P(C), 0.21190. As a finding, since the ASSISTment iBKT with P(C) shows a low RMSE score, it is considered to have a higher degree of prediction accuracy. RMSE result for dataset KDD iBKT without P(C) shows 0.2826, which is higher than KDD iBKT with P(C), 0.2590. Again, the result reflects the same finding: a higher degree of prediction accuracy where the RMSE for dataset KDD iBKT with P(C) shows the lower RMSE score.

Second, we compare the RMSE differences based on specific iBKT, without P(C) and with P(C). The result for RMSE value differences for both datasets (Fig. 6) without P(C) is 0.0507. However, the result for RMSE value differences for both datasets

Fig. 6. RMSE value for dataset ASSISTment iBKT without P(C) and KDD iBKT without P(C).

the value as NA. The final value to be evaluated is the RMSE result without and with P(C).

Using the final RMSE value, we plot the bar graph to see the comparison. The bar graph in Fig. 6 depicts the RMSE value for

Table V.   Details analysis of comparison

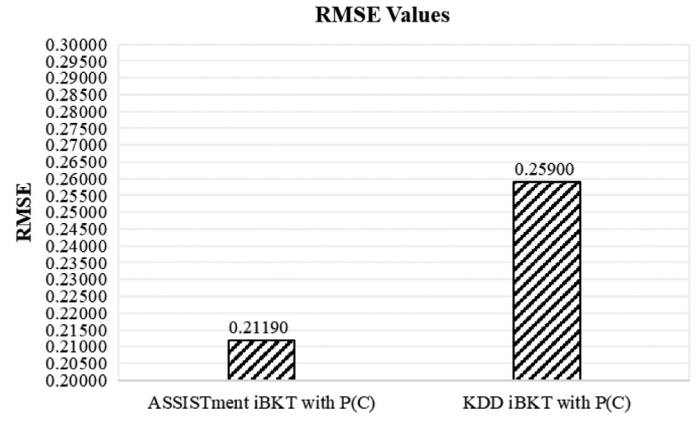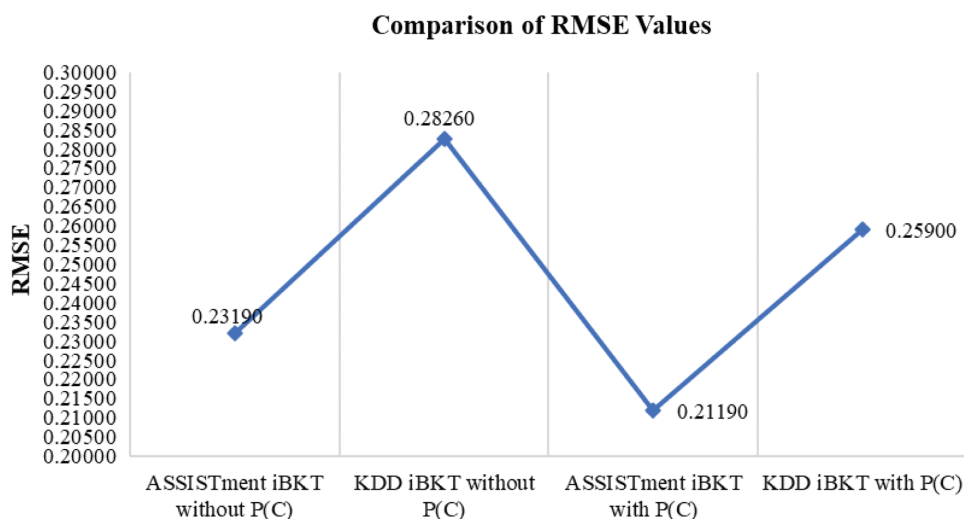| Comparison	Findings
Specific individual dataset, ASSISTment and KDD	• ASSISTment iBKT:
	○ RMSE dataset ASSISTment iBKT without P(C) is 0.23190.
	○ RMSE dataset ASSISTment iBKT with P(C) is 0.21190.
	○ Conclusion: ASSISTment iBKT with P(C) shows a low RMSE score. Thus, it is considered to have a higher degree of prediction accuracy.
	• KDD iBKT:
	○ RMSE dataset KDD iBKT without P(C) is 0.28260.
	○ RMSE dataset KDD iBKT with P(C) is 0.25900.
	○ Conclusion: KDD iBKT with P(C) shows a low RMSE score. Thus, it is considered to have a higher degree of prediction accuracy.
Specific iBKT, without the new parameter P(C) and with the new parameter P(C)	• RMSE value difference in Fig. 6 without P(C) is 0.0507.
	• RMSE value differences for both datasets with P(C) shows 0.04710.
	• Conclusion: The RMSE with confidence parameter P(C) clearly shows which dataset can provide higher prediction accuracy. The ASSISTment dataset provides a higher prediction when applying the new parameter, which is 0.21190

Comparison of RMSE Values

Fig. 8. Comparison of RMSE values for dataset ASSISTment iBKT and KDD iBKT without and with P(C).

(Fig. 7) with P(C) shows 0.04710. Based on the findings, we conclude that the proposed P(C) solution clearly shows which dataset can provide higher prediction accuracy. In this study, the ASSISTment dataset provides a higher prediction when applying the new parameter, which is 0.21190. Therefore, this will affect the value of iBKT from both datasets when implementing each value of parameter received to the formula of iBKT.

To ease the visualization, we plot the line graph combining the result as in Fig. 8 to compare the RMSE values of the iBKT without and with P(C). The RMSE value of iBKT with P(C) has a narrower range value than iBKT from both datasets, ASSISTment and KDD. As a result, it is concluded that utilizing iBKT with a confidence level to predict student performance is more accurate and reliable than using standard iBKT without P(C).

V. CONCLUSION

In conclusion, this research analyzed that ITSs are computer systems that aim to improve students' performance in learning by providing personalized instruction and feedback to the user using AI technology. ITS contains four modules, and this research only focuses on one module, the student's module, where the prediction and measurement of the student's performance happened. Predicting student performance in ITS has always been a highlighted issue stated by other researchers. This paper analyzed the available methods to predict students' performance. We found that iBKT can fill the gap regarding students' performance prediction. Then, we proposed a new prediction approach by adding a parameter regarding confidence, P(C), to iBKT. We compare the performance accuracy of the iBKT and the enhanced iBKT with the confidence parameter using ASSISTment and KDD datasets. It is observed that since the value of RMSE is lower, the iBKT method with P(C) is more accurate. The ASSISTment dataset shows a low RMSE value compared to KDD. With the result, it is proven that the proposed prediction approach is more reliable and precise. For future studies, we plan to explore ITS predictions apart from student performance in student modules and also to consider other parameters aside from the confidence level to deal with the student uncertainty problem in measuring and predicting the students' performance in ITS.

AUTHOR CONTRIBUTIONS

Almarzuki HF performed data collection and processing, and writing – review and editing; Samah KAFA performed model analysis on the equation, writing – review, final editing, and funding acquisition; Rahim SKNA performed the literature review, data analysis, and writing – review and editing; Ibrahim S performed analysis, visualization and graphic improvement, and final editing; and Riza LS performed methodology and visualization and graphic improvement, and writing – review and editing. All authors had approved the final version.

REFERENCES

[1] N. K. Elfaki, I. Abdulraheem, and R. Abdulrahim, "Impact of e-learning vs traditional learning on students' performance and attitude," *Int. J. Med. Res. Health Sci.*, vol. 8, no. 10, pp. 76–82, 2019.

[2] K. Crowston et al., "Knowledge tracing to model learning in online citizen science projects," *IEEE Trans. Learn. Technol.*, vol. 13, no. 1, pp. 123–134, 2019. DOI: 10.1109/TLT.2019.2936480.

[3] I. Subchi, M. Suparta, Zulkifli, S. U. Masruroh, N. F. Rozi, and N. S. Nuraeni, "The development of e-learning with intelligent tutoring system through anthropological approach (a case study at Uin Syarif Hidayatullah Jakarta)," 2019. DOI: 10.1109/CITSM47753.2019.8965357.

[4] H. M. Sh Bakeer and S. S. Abu-Naser, "An intelligent tutoring system for learning TOEFL," *Int. J. Acad. Pedagogical Res.*, vol. 2, no. 12, pp. 9–15, 2018.

[5] M. Laaziri, K. Benmoussa, S. Khoulji, K. Mohamed Larbi, and A. El Yamami, "Implementation of an intelligent tutoring system for the use of university governance information systems," in *SCA '18: Proc. 3rd Int. Conf. Smart City Appl.*, 2018, pp. 1–7. DOI: 10.1145/3286606.3286819.

[6] T. Käser, N. R. Hallinen, and D. L. Schwartz, "Modeling exploration strategies to predict student performance within a learning environment and beyond," in *LAK '17: Proc. Seventh Int. Learn. Anal. Knowl. Conf.*, 2017, pp. 31–40. DOI: 10.1145/3027385.3027422.

[7] S. Klingler, A. G. Schwing, and M. Gross, "Dynamic Bayesian networks for student modeling," *IEEE Trans. Learn. Technol.*, vol. 10, no. 4, pp. 450–462, 2017. DOI: 10.1109/TLT.2017.2689017.

[8] Z. Wang, J. Zhu, X. Li, Z. Hu, and M. Zhang, "Structured knowledge tracing models for student assessment on Coursera," in *L@S 2016 – Proc. 3rd 2016 ACM Conf. Learn. Scale*, 2016, pp. 209–212. DOI: 10.1145/2876034.2893416.

[9] B. Lewis, I. Smith, M. Fowler, and J. Licato, "The robot mafia: A test environment for deceptive robots," in *28th Modern Artif. Intell. Cognit. Sci. Conf., MAICS 2017*, 2017, pp. 189–190. DOI: 10.1145/1235.

[10] N. Thai-nghe, "Predicting student performance in an intelligent tutoring system," PhD, Stiftung Universität Hildesheim, 2011.

[11] M. V. Yudelson, K. R. Koedinger, and G. J. Gordon, "Individualized Bayesian knowledge tracing models," *Lect. Notes Comput. Sci. (Incl. Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinf.)*, vol. 7926 LNAI, pp. 171–180, 2013. DOI: 10.1007/978-3-642-39112-5-18.

[12] T. Schodde, K. Bergmann, and S. Kopp, "Adaptive robot language tutoring based on Bayesian knowledge tracing and predictive decision-making," in *Proc. 2017 ACM/IEEE Int. Conf. Human-Robot Interact.*, 2017, pp. 128–136. DOI: 10.1145/2909824.3020222.

[13] P. Dašić, J. Dašić, B. Crvenković, and V. Šerifi, "A review of intelligent tutoring systems in e-learning," *Ann. Oradea Univ. – Fascicle Manag. Technol. Eng.*, vol. 15, 2017, pp. 85–90, 2016. DOI: 10.15660/AUOFMTE.2016-3.3276.

[14] H. A. Al Rekhawi, "Android applications development intelligent tutoring system," *Int. J. Acad. Inf. Syst. Res.*, vol. 4, no. 5, pp. 33–58, 2020.

[15] H. A. A. Hasanein and S. S. Abu-Naser, "Developing education in Israa University using intelligent tutoring system," *Int. J. Acad. Pedagogical Res.*, vol. 2, no. 5, pp. 1–16, 2018.

[16] M. A. Hamed, S. S. Abu-Naser, and K. S. Abualhin, "Intelligent tutoring system effectiveness for water knowledge and awareness," *Int. J. Acad. Inf. Syst. Res.*, vol. 2, no. 4, pp. 18–34, 2018.

[17] M. U. Modén, M. Tallvid, J. Lundin, and B. Lindström, "Intelligent tutoring systems: why teachers abandoned a technology aimed at automating teaching processes," in *Proc. Annu. Hawaii Int. Conf. Syst. Sci.*, 2021, pp. 1538–1547. DOI: 10.24251/hicss.2021.186.

[18] B. Chakraborty and M. Sinha, "Student evaluation model using Bayesian network in an intelligent e-learning system," *J. Inst. Integr. Omics Appl. Biotechnol.*, vol. 7, no. 2, pp. 51–60, 2016.

[19] S. Minn, M. C. Drsmarais, F. Zhu, J. Xiao, and J. Wang, "Dynamic student classification on memory networks for knowledge tracing," in *Adv. Knowl. Disc. Data Mining: PAKDD 2019*, 2019, pp. 163–174.

[20] M. Haridas, N. Vasudevan, S. Gayathry, G. Gutjahr, R. Raman, and P. Nedungadi, "Feature-aware knowledge tracing for generation of concept-knowledge reports in an intelligent tutoring system," in *Proc. IEEE Int. Conf. Technol. Educ., T4E 2019*, 2019, pp. 142–145. DOI: 10.1109/T4E.2019.00-34.

[21] P. Nedungadi and M. S. Remya, "Incorporating forgetting in the personalized, clustered, Bayesian knowledge tracing (PC-BKT) model," 2015. DOI: 10.1109/CCIP.2015.7100688.

[22] S. Doroudi and E. Brunskill, "The misidentified identifiability problem of Bayesian knowledge tracing," in *Proc. 10th Int. Conf. Educ. Data Mining, EDM 2017*, Institute of Education Sciences (ED), 2017, pp. 143–149.

[23] Z. MacHardy and Z. A. Pardos, "Toward the evaluation of educational videos using Bayesian knowledge tracing and big data," in *L@S 2015 – 2nd ACM Conf. Learn. Scale*, 2015, pp. 347–350. DOI: 10.1145/2724660.2728690.

[24] Z. A. Pardos and N. T. Heffernan, "Modeling individualization in a Bayesian networks implementation of knowledge tracing," *Lect. Notes Comput. Sci. (Incl. Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinf.)*, vol. 6075 LNCS, pp. 255–266, 2010. DOI: 10.1007/978-3-642-13470-8_24.

[25] Y. Wang, "Student modeling within a computer tutor for mathematics: using Bayesian networks and tabling methods," PhD, Worcester Polytechnic Institute, 2015.

[26] P. I. Pavlik, H. Cen, and K. R. Koedinger, "Performance factors analysis—a new alternative to knowledge tracing," in *Proc. 14th Int. Conf. on Artificial Intelligence in Education*, Institute of Education Sciences (ED), 2009, pp. 1–8.

[27] Y. Gong, J. E. Beck, and N. T. Heffernan, "Comparing and optimizing knowledge tracing and performance factor analysis," *Int. J. Artif. Intell. Educ. Accept.*, vol. 21, no. 1–2, pp. 27–46, 2011.

[28] Y. Gong, J. E. Beck, and N. T. Heffernan, "Comparing knowledge tracing and performance factor analysis by using multiple model fitting procedures," *ITS2010 Intell. Tutor. Syst. LNCS*, vol. 6094, pp. 35–44, 2010. DOI: 10.1007/978-3-642-13388-6_8.

[29] M. V. Yudelson, "Individualizing Bayesian knowledge tracing. Are skill parameters more important than student parameters?," in *Proc. 9th Int. Conf. Educ. Data Mining, EDM 2016*, Institute of Education Sciences (ED), 2016, pp. 556–561.

[30] Y. Wang and N. Heffernan, "A comparison of two different methods to individualize students and skills," *Lect. Notes Comput. Sci. (Incl. Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinf.)*, vol. 7926 LNAI, pp. 836–839, 2013. DOI: 10.1007/978-3-642-39112-5-125.

[31] P. Yadav, "Comparative analysis of intelligent tutoring system approaches," *J. 4 Res.*, vol. 03, no. 10, pp. 7–10, 2017.

[32] Y. Gong, "Student modelling in intelligent tutoring systems," *Artif. Intell. Rev.*, vol. 7, no. 3–4, pp. 227–240, 2014. DOI: 10.1007/BF00849556.

[33] S. K. N. A. Rahim, A. H. Jaafar, S. Masrom, F. Zulkipli, S. R. Ismail, and N. Ahmad, "Intelligent tutoring systems' measurement and prediction of students' performance using predictive function," *Int. J. Emerg. Trends Eng. Res.*, vol. 8, no. 1 1.1 Special Issue, pp. 187–192, 2020. DOI: 10.30534/ijeter/2020/2981.12020.

[34] N. T. Heffernan, "2015 ASSISTments skill builder data," 2015. https://sites.google.com/site/assistmentsdata/datasets/2015-assistments-skill-builder-data?authuser=0

[35] J. Stamper, A. Niculescu-Mizil, S. Ritter, G. J. Gordon, and K. R. Koedinger, "Challenge data set from KDD Cup 2010 educational data mining challenge," 2010. [Online]. Available http://pslcdatashop.web.cmu.edu/KDDCup/downloads.jsp

[36] R. S. J. D. Baker, Z. A. Pardos, S. M. Gowda, B. B. Nooraei, and N. T. Heffernan, "Ensembling predictions of student knowledge within intelligent tutoring systems," *Lect. Notes Comput. Sci. (Incl. Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinf.)*, vol. 6787 LNCS, pp. 13–24, 2011. DOI: 10.1007/978-3-642-22362-4_2.

[37] R. Pelánek, "Metrics for evaluation of student models," *J. Educ. Data Min.*, vol. 7, no. 2, pp. 1–19, 2015.