Rodrigo, M. M. T. et al. (Eds.) (2021). Proceedings of the 29th International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

A Machine Learning Approach to Estimating Student Mastery by Predicting Feedback Request and Solving Time in Online Learning System

Kannan N[a*], Charles Y. C. YEH[a], Chih-Yueh CHOU[b] & Tak-Wai CHAN[a]
[a]*Graduate Institute Of Network Learning Technology, National Central University, Taiwan*
[b]*Department of Computer Science and Engineering, Yuan Ze University, Taoyuan, Taiwan*
*kannannataraj@g.ncu.edu.tw

Abstract: One of the most significant challenges for computers in education is the capacity to provide intelligent and adaptable learning systems to meet the real needs of students. In order to create efficient adaptive or personalized mechanisms for educational content, student models are proposed to estimate the actual knowledge or mastery level of students. Some earlier student models were proposed to estimate student mastery based on the correctness (e.g., correct or incorrect) of responses, feedback request, and solving time using classical Markov process and logistic regression models. In particular, these models were applied to predicting student future correctness, feedback request, and solving time (i.e., on the next question).The advent of increasingly large-scale datasets has turned Machine Learning (ML) methods such as conventional machine-learning algorithms and deep learning models for prediction into competitive alternatives to classical Markov process and logistic regression models. In addition, prediction by ML methods has numerous advantages such as interpretability, good accuracy, ease of maintenance, less execution time, and appropriately handling of missing data. Moreover, recent studies exhibit the significant achievement of ML prediction methods for estimating students' performance and mastery using learning log data (i.e., correctness, feedback request level, solving time, etc.). Hence, it is reasonable to use ML methods to estimate student mastery by predicting the feedback request level and solving time. This study analyzed the data logged by an online learning system called Math-Island, which teaches elementary level mathematics by incorporating game mechanisms and scaffolding feedback. Machine-learning regression methods such as Multiple Linear Regression (MLR), Support Vector Regression (SVR), Random Forest Regression (RFR), Extra Trees (ET), and Gradient Boosting Regression (GBR) were applied. The results showed that RFR and GBR were found to outperform other models to predict future feedback request level and solving time. The results lead to several future works. First, incorporating ML predictive models into Math-Island tutoring system to identify the individual student's actual needs and; reduce learning loss substantially. Second, it drives to effectively build a more efficient adaptive mechanism within the current session to utilize students' active learning time.

Keywords: Machine learning, prediction, regression, online learning system.

1. Introduction

Estimating students' learning performance and content knowledge or mastery is a long-standing practice in the education field, which contributes to supporting and enhancing learning achievements. Mastery-oriented goals focus students' attention on achievement based on intrapersonal learning standards; performance goals focus on achievement based on normative or comparative performance standards. Ansems et al. (2019) describes that mastery goals focus on developing competence and mastering a task. In contrast, performance goals focus on the demonstration of competence and outperforming others. Research evidence suggests a mastery goal orientation that promotes a motivational pattern likely to promote long-term and high-quality involvement in learning (Tuominen, Juntunen, & Niemivirta, 2020). Moreover, studies show that mastery-oriented goals consistently lead to intrinsically motivated, self-regulated learning and promote comprehension (Ansems et al., 2019; Caniëls, Chiocchio, &vanLoon, 2019). On the other hand, researchers suggested one-to-one technology (Chan et al., 2006) through which every student is equipped with a device to learn in school or at home seamlessly. Online learning systems (Jin, 2020; Yeh, Cheng, Chen, Liao, & Chan, 2019) successfully encompass this feature: they teach skills, such as algebra, numeric operation, geometry, computer programming, or medical diagnosis, using mastery-oriented goal principles and provide learners with individualized feedback and materials adapted to their level of understanding (Romero, Hernández, Juola, Casadevante, & Santacreu, 2020). Studies have demonstrated that online learning has gained much attention in recent years. However, it needs improvements and incorporates new technologies as per large-scale datasets (Ogdanova, Tova, & Vetaeva, 2014) to make an efficient adaptive and flexible learning context.

Mastery estimation and prediction is used to identify what a student will do or know at the end of an instructional unit (Bälter, Zimmaro, & Thille, 2018). Mastery estimation has been varying in line with the evolution of tutoring and learning methods. The well-known student models, Bayesian knowledge tracing (BKT) by Corbett and Anderson (1994), and performance factor analysis (PFA) by (Pavlik, Cen, & Koedinger, 2009), have been widely used to estimate current student knowledge mastery. BKT was proposed during the 1990s; BKT predicts student mastery via probabilities that include four parameters per knowledge component. PFA uses logistic regression to predict mastery as the output of the learned or unlearned state. These models predict student mastery based on the correctness of responses. In particular, these models attempt to predict students' future correctness (i.e., on the next question). However, there are few significant limitations to these methods; for example, BKT assumes no downsides and cannot forget once students learn the skill. In contrast, Bälter (2018) mentioned that once learned, there are also downsides with continued repetition and might even be detrimental under certain conditions. Moreover, previously reported results to show that BKT cannot handle missing data patterns for prediction (Gervet, Koedinger, Schneider, & Mitchell, 2020). Furthermore, Pavlik et al. (2009) implemented the PFA algorithm in Excel, and using Excel is not feasible for large-scale datasets and on-time supports. Moreover, both models were not fully able to account for more rapid shifts in student performance, especially in cases where a student struggles early but goes on to drastically improve their performance (Slater & Baker, 2019). Therefore, these limitations have turned machine-learning methods for prediction into competitive alternatives to classical Markov process and logistic regression models. Furthermore, Machine-learning methods such as Random Forest, linear regression, feed-forward neural network, and deep learning models such as Deep Knowledge Tracing were all shown to deliver superior results to BKT and PFA on their own (Mao, Lin, & Chi, 2018; Piech et al., 2015).

Machine-Learning (ML) methods can be applied throughout science, technology, and commerce, leading to more evidence-based decision-making across various fields, including education, health care, manufacturing, financial modeling, policing, and marketing. It mainly has a considerable impact on the educational field, especially for estimating, tracing and predicting students' mastery and performance (Imran, Latif, Mehmood, & Shah, 2019; Sokkhey & Okazaki, 2020). In the technology-pervasive world, as a student is working toward a solution, the system keeps track of his or her actions and provides feedback to help the student progress (Myneni, Narayanan, Rebello, Rouinfar, & Pumtambekar, 2013). A study (Lai & Lin, 2015) found that students received immediate, elaborative, text-based feedback led to more effective learning and higher motivation. Results suggest that immediate feedback from computer-based learning tasks benefit both high and low prior knowledge students, with low prior knowledge students exhibiting more significant gains (Razzaq, Ostrow, & Heffernan, 2020). Solving time is based on correctness, which is collected separately within each attempt. Student solving time has been mainly used to assess student learning because it can indicate how active and accessible student knowledge is (Mao et al., 2018). For example, it has been shown that solving time reveals student mastery, and has been suggested as an indicator of student engagement in answering questions as well as an important factor for predicting motivation in an e-learning environment (Schnipke, 2002).

Therefore, in this work, we used both feedback request level and solving time as feature variables for prediction using ML methods to estimate students' hidden mastery level. ML methods thrive on large datasets, create assumptions about the data, and allow them to use non-normally distributed input variables. Moreover, the ML approach has several advantages: interpretability, accuracy, ease of maintenance, adequate execution time, and appropriately handling missing data, unlike BKT, PFA. In addition, we evaluate models only in terms of the accuracy of their predictions—the resulting best models to employ for carrying out adaptive pedagogical strategies and further advanced personalized learning for future works.

2. Machine Learning Methods in Student Performance Prediction

Students' learning progress often refers to student knowledge as a latent variable (Corbett &Anderson, 1994). Estimating actual knowledge of the content and providing enough practice opportunities before moving on to the next level are challenging for educators, particularly in an online learning environment or distance learning and flexible learning context. Pelánek (2015) mentioned that the over-practice (the practice of items that the student already mastered, i.e., "wasted time" of students) and under-practice (a missing practice that is necessary for mastery of a topic and further progress) are the broader problems for learners, which is driven them to loss of learning. In a study of high school students participating in the Optimized Cognitive Tutor geometry curriculum, it was found that 58% out of 4102 practices and 31% of 636 exercise questions were done after the students had reached mastery (Cen, Koedinger, & Junker, 2007). However, it is possible to reduce study time without the loss of learning (Bälter et al., 2018). A prior study conducted with 265 individuals shows that over half the individuals are expected to need less than five practice opportunities to reach mastery. In addition, over 40 students require at least 15 practice opportunities, and over 30 of those require 20 or more to reach mastery (Lee & Brunskill, 2012). Moreover, Beck and Gong (2013) found that 69% of students in Cognitive Algebra Tutor (CAT) and 62% of students in ASSISTments have mastered the skill after ten practice opportunities.

The findings above suggest that the optimal number of questions needed to prevent students from running out of insufficient practice opportunities before they have mastered a skill, simultaneously not spending more time and resources on an already mastered skill (Bälter et al., 2018). Slater &Baker (2019) recommend that mastery or knowledge estimation is a valuable technique for quickly identifying student's needs likely to wheel-spin (Beck & Gong, 2013) and providing additional scaff olding or support for their learning. Therefore, estimating student mastery of skill after solving the optimal number of questions could alleviate the loss of learning for high-low achieving students.

Machine learning is widely used for prediction problems, especially in the education field. In other words, machine-learning methods showed good performance than other statistical methods. For instance, recent studies have demonstrated good accuracy in predicting student performance in contexts of solving problems in Intelligent Tutoring Systems (ITSs) or completing courses in a classroom or in Massive Open Online Courses (MOOC) platforms (Jin, 2020; Mao et al., 2018). In most of the previous literature (Table 1), many researchers have approached to find the best algorithms for estimating future performance based on predicting correctness, final grade, dropout status, and final exam scores; moreover, most studies are proposed classification—few used regression methods due to the nature and relationship of each input variable with the output variable. In addition, the studies mentioned earlier have used longer timescale outcome prediction due to the dataset and variables (i.e., semester result, score, and grade). Different models have different sensitivities to the type of predictors in the model; how the predictors enter the model is also important.

Table 1. *Summary of Recently Proposed ML Methods in ITSs for Predict Student Performance*

Algorithm	Reference	Study focus	Input Variables	Output Variable
Naïve Bayes, Generalized Linear Model, Logistic Regression, Deep Learning, Decision Tree, Random Forest, and Gradient Boosted Trees.	Abidi, Hussain, Xu, &Zhang, 2018	Academic performance of students	Attempt_count, hint_total, overlap_time, response type (correct or incorrect)	Final grade
Artificial Neural Networks, Support Vector Machines, Logistic Regression, Naïve	Hussain, Zhu, Zhang,	Student's difficulty during the next session	Average time, total number of activities, average idle time, average number of	Grades

bayes Classifiers and Decision trees	Abidi, &Ali, 2019		keystrokes and total related activity	
Logistic Regression	Asselman, Khaldi, &Aammou, 2020	Prior required scaffolding items to predict future student performance	Prior scaffolding	Correctness
Multilayer Perceptrons , Sequential Minimal Optimization of Support Vector Machine, Logistic Regression, and Random Forest.	Sokkhey &Okazaki, 2019	Student performance in mathematics	Domestic, academic, and attitudinal information describing each student,	Grade
Linear Regression	Rong &Bao-Wen, 2018	Students at risk of failure	Prior term's student clicker data and final exam scores, high school Grade Point Average (GPA), high school GPA, gender, age, socioeconomic status	Final exam score
Support Vector Machine, Random Forest, and Extra Trees	Jin, 2020	Students' early dropout status in MOOC.	Clickstream data	Drop out status

Accordingly, this study applies multiple regression algorithms such as Multiple Linear Regression (MLR), Support Vector Regression (SVR), Random Forest Regression (RFR), Extra Trees (ET), and Gradient Boosting Regression (GBR) to predict numerical output variables (i.e., feedback request level and solving time) to estimating student mastery, which happens in a specific time (i.e., shorter timescale). Moreover, in this paper, we evaluate models only in terms of the accuracy of their predictions.

3. Method

This study's data were collected from students in Taiwan who used the "Math-Island online learning game" during 2016-2019. Math-Island is an online learning system that incorporates gamified knowledge map of the elementary mathematics curriculum (Yeh et al., 2019). The Math-Island game targets the mathematics curriculum of elementary schools in Taiwan, mainly containing the four domains: numerical operation, quantity and measure, geometry, and statistics and probability (Figure 1 a). Each domain contains gamified knowledge map of concept units for students to learn mathematics by units (Figure 1 b). Every unit has a different quantity of tasks (Figure 1 c), and each task has several problem-solving questions (Figure 1 d). Students can freely choose the learning path according to their interests, and if students face difficulty while learning the skill, the system will provide scaffolding feedback.

The system collects many data when students answer problem-solving questions, including problem-related information, and instructional assistance, e.g., student id, number and level of feedback request, solving time, and total log time on individual questions. We use Pearson's correlation coefficient (Figure 3) method to choose feature variables; the plot showing the positive correlation coefficient between the input variables and output variable, which determined that the regression algorithms are more appropriate for prediction. Therefore, this study applied machine-learning regression methods to predict feedback request level and solving time using the following variables:

Feedback-level is the total number of feedback or attempts students used each question (minimum 0 - maximum 3). (0) Feedback (such as "correct"), (1) Feedback (such as "incorrect"), (2) Feedback + scaffolded solution (how to answer), (3) All three services with the answer (Figure 2).

Solving time is how long it takes for students to finish each question, excluding every feedbacks time if students answered the questions incorrectly. We used these variables to build two different prediction models accordingly. Variables collected from the first four questions are applied to predict the variable of the fifth question (Table 2).

Table 2. *Description of Input Variables and Output Variable*

Independent variables				Dependent variable
Question 1 x_1	Question 2 x_2	Question 3 x_3	Question 4 x_4	Question 5 Y
Feedback request level	Feedback request level	Feedback request level	Feedback request level	Feedback request level
Solving time	Solving time	Solving time	Solving time	Solving time

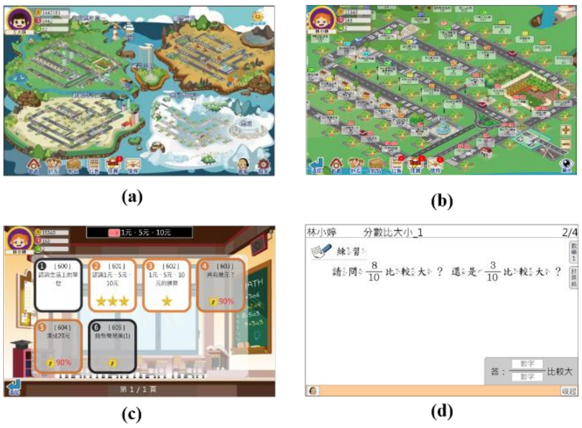

Figure 1. Math-Island online learning system and task screen

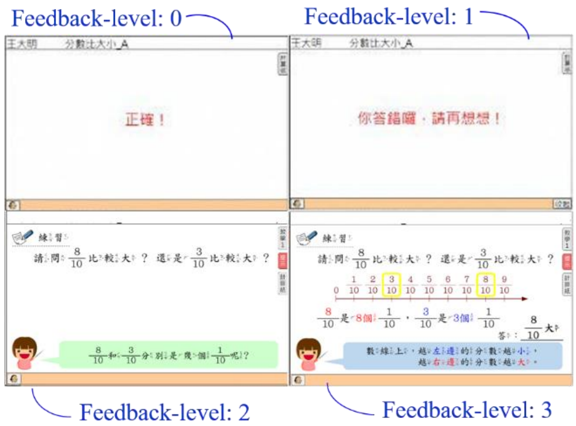

Figure 2. Feedback-levels

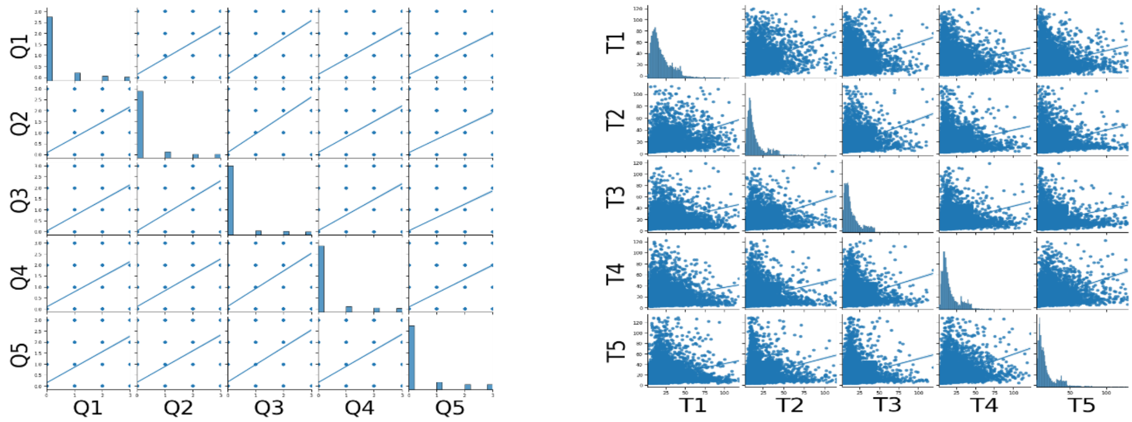

Figure 1. Relationship between Input variables and Output Variables. (Pair plot distribution shows the positive correlation coefficient between the input variables and output variable)
Keys: Q1-Q5 – feedback levels of Question 1 to Question 5; T1-T5 – Solving time of Question 1 – Question 5

Our approach is to estimate students' mastery by predicting their feedback request level and solving time. We assume that the predictive output of feedback request level (0-1) and solving time (5-15) is the threshold of mastery (Figure 4); nevertheless, more exploratory study is needed to establish our findings here. However, this study aims to estimate the regression model with coefficients of c, w0 + w1 x1 + w2 x2 + … + wn and fit the training data with minimal squared error and predict the output y for each to find the best prediction model.

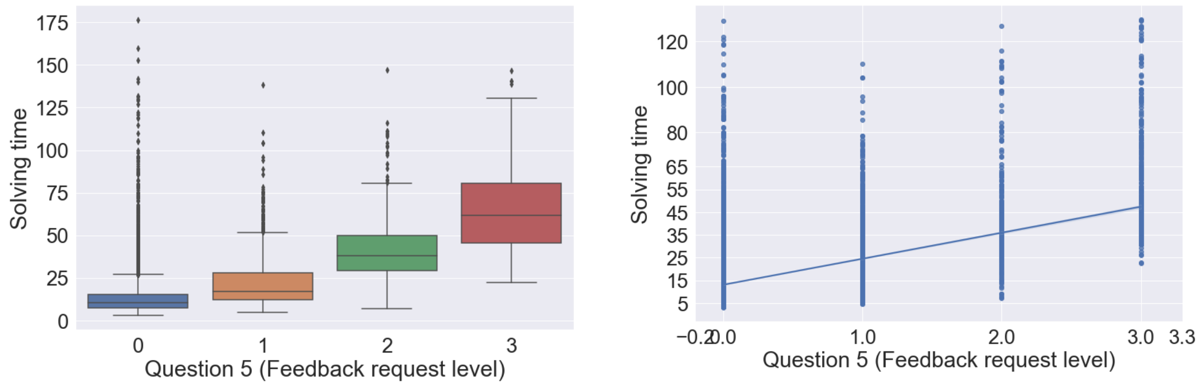

Figure 2. Distribution of utilized feedback request level according to the solving time on question five (left: the box plot distribution of feedback request level according to solving time, right: the scatterplot of two variables with 95% confidence interval of regression line)

4. Data Collection and Analysis

For the initial stage of this study, we used data collected from one task unit in the numerical operation domain in Math-Island, where there are 50 task units, and the level of difficulty is identical for all questions. The data contains 13168 answer records, including student id, mission id, feedback request level, and solving time of each question. After removing outliers (e.g., less than 1.5 sec, more than 140 sec, and null values), 12034 records were chosen for this study.

Table 3. *Dataset Description.*

Sl.no	Q1	Q2	Q3	Q4	Q5	T1	T2	T3	T4	T5
0	3	1	0	0	3	62.811	35.696	10.104	10.527	79.831
1	3	0	0	0	0	33.524	9.343	11.446	11.017	11.524
2	3	0	0	0	1	69.260	13.187	25.633	18.066	33.733
3	0	0	0	0	0	9.065	10.573	26.696	8.785	5.783
…	...	...	...	...	...	...	...	...	...	...
12033	2	2	3	2	2	35.000	39.000	35.000	35.000	45.000
min	0	0	0	0	0	1.68	2.80	2.56	2.61	3.17
max	3	3	3	3	3	119.23	113.91	119.38	122.46	129.66
count	12034	12034	12034	12034	12034	12034	12034.	12034	12034	12034
mean	0.363	0.322	0.276	0.329	0.394	20.27	16.103	16.031	17.817	17.562
SD	0.79	0.78	0.74	0.79	0.84	15.14	12.67	12.24	13.94	15.08

5. Model Evaluation

The dataset has split into a training dataset and test dataset within 75-25 ratio. We use two standard regression metrics Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE), to evaluate applied methods' performance. The evaluation included a comparison between the prediction results of the models (Table 4). Note that both of these metrics measure error, so a smaller number indicates a better predictive model. [1]

The simple baseline was measured by central tendency measurement that used the global mean of Y and then calculated MAE and RMSE of the mean by reducing the y_test data. The baseline was measured to infer the performance of each model.

[1] This study used Python 3.6 to analyze the data with the following kits: the Numpy suite for data collation, Pandas, Matplotlib, the SciPy kit for data visualization, Seaborn, and the Scikit-learn kit.

Table 4. *Model Evaluation and Comparison between the Models*

Algorithms	Feedback request level		Solving time	
	MAE	RMSE	MAE	RMSE
Baseline	**0.606**	**0.827**	**10.418**	**14.906**
Multiple Linear Regression	0.335	0.543	7.628	12.856
Support Vector Regression	0.384	0.550	8.065	14.763
Random Forest Regression	0.333	0.530	7.361	12.756
Extra Trees Regression	0.335	0.530	7.461	12.894
Gradient Boosting Regression	0.352	0.529	7.030	12.342

Table 5. *Comparison of Actual (y_test) and Prediction (y_pred) Values*

	Feedback request level (y_pred)						Solving time (y_pred)					
	y_test	MLR	SVR	RFR	ET	GBR	y_test	MLR	SVR	RFR	ET	GBR
count	3009	3009	3009	3009	3009	3009	3009	3009	3009	3009	3009	3009
mean	**0.37**	0.39	0.30	**0.38**	**0.38**	**0.38**	**17.14**	**17.37**	12.96	17.86	18.00	**17.32**
SD	0.82	0.62	0.66	0.64	0.64	0.63	14.90	7.86	3.41	9.67	9.63	8.61
min	0	0.10	**-0.08**	0.00	0.00	0.10	3.33	7.01	3.67	4.26	4.37	4.29
max	3	2.84	2.90	2.78	3.00	2.52	126.83	59.84	23.37	91.06	83.54	63.91

Table 4 shows the performance comparison among each model to predict the final question's feedback request level and solving time. Table 5 shows the comparison of each model's prediction outputs with the actual value. The result of the five models has outperformed baseline MAE and RMSE. However, with comparing each model for feedback request level, RFR (0.333), MLR (0.335), and ET (0.335) seem to have minimal MAE than other models; on the other hand, GBR has minimal RMSE (0.529). In addition, for solving time, GBR (MAE: 7.030, RMSE: 12.342) outperformed other models.

MLR (MAE: 0.335, RMSE: 0.543) for feedback request level prediction performed better than baseline, and it produced the closest prediction output with the actual value; standard deviation (SD) also is significantly less than other predictors. On the other hand, for solving time, MLR (MAE: 7.628, RMSE: 12.856) performed better than RFR, ET, SVR. However, the prediction of MLR is (Min; 7.865, Max 59.844) significantly lesser than the actual (Min: 3.17, Max: 129.66) value, including RFR and ET.

SVR for predicting feedback request level (MAE: 0.384, RMSE: 0.550) and solving time (MAE: 8.065, RMSE: 14.763) has significantly performed worse than all other models. Furthermore, Table 5 shows SVR has an exploration issue on the prediction value (-0.081) for the feedback request level.

RFR for predicting feedback request level (MAE: 0.333, RMSE: 0.530) and solving time (MAE: 7.361, RMSE: 12.756), ET for predicting feedback request level (MAE: 0.335, RMSE: 0.530) and solving time (MAE: 7.461, RMSE: 12.894) had similar performance than MLR and SVR because ensemble models can be performed well on non-linear data. Moreover, the feedback request level's prediction output in Table 5 shows significantly similar with actual value (RFR: Min: 0.00 - Max: 2.780, ET: Min: 0.000 - Max: 3.000).

GBR for predicting feedback request level (MAE: 0.352, RMSE: 0.529) and solving time (MAE: 7.030, RMSE: 12.342) outperforms all other models. Moreover, for predicting solving time, the mean of GBR's prediction (17.327) is almost similar to the mean of MLR's prediction Mean (17.370).

In our analysis, every model has performed almost similarly shown in Table 4. However, for feedback request level, ensemble methods of RFR, ET; for solving time, GBR and MLR were found better than other models. One potential explanation is that the ensemble methods enrich the skills and generate more outcomes than expert linear models, which may also be a better model for the domain.

6. Discussion and Future Work

The study aims to explore a better method to estimate the mastery of math knowledge at the elementary level, which is taught by learning-by-units based on mastery-oriented goal principles. Our approach was to estimate students' mastery by predicting their feedback request level and solving time. The prediction model was applied to assess student's work on each task. The results revealed the plausibility of predicting students' feedback request level and solving time. Popular machine learning regression methods are used due to their broad adoption within the field of educational data mining and learning analytics and its' relative computational simplicity. The ensemble model of RFR and ET performed better than other models to predict future feedback request level; on the other hand, MLR and GBR outperform other models for the prediction of solving time.

The application of machine learning addresses how to build online learning systems that improve automatically through experience to assess student mastery and provide additional support. Prediction models used the log data of student's feedback request level and solving time of four questions to predict the student's feedback request level and solving time of the next question. Students are at various levels, have different skills and different learning abilities. Making high-quality predictions immediately available may help instructors, teaching assistants and intelligent learning systems identify groups of students who are wheel-spinning and need alternate forms of assistance (Slater & Baker, 2019). Without support, a student who is not making progress in an open and distance learning context is likely to drop out of the course (Tang, Xing, & Pei, 2018). If we can identify this struggle before it goes on for too long, we may be able to address the problem and help students get back on track. Researchers (Tuominen et al., 2020; Yu et al., 2018) have been done on recognizing the student's affective state and responding to it actively. The prediction model of this study can be modified for predicting students' affective state while the student's affective state data is collected.

Based on the result, the use of ML methods deserves more attention on the prediction of mastery or knowledge in future studies. Furthermore, this study encourages researchers interested in using ML methods and other cutting-edge knowledge tracing algorithms to predict mastery of the skill and make the path to move further direction (i.e., additional support or move on the next level), rather than just predicting performance within-data. Bälter (2018) described that OLI courses do not provide students with a path for moving on after mastering a skill, either through a forced pathway or by providing data to the student that he/she had mastered the skill. For adaptive behavior, accurate predictions matter most, while for actionable insights, the interpretability and the stability of parameter estimates supersede accuracy; open and distance learning learner models require both.

Overall, we make the following contributions: 1) our work makes the initial process of estimating students' mastery by predicting their feedback request level and solving time on the next question for developing an efficient adaptive mechanism. 2) We explored the robustness and effectiveness of the proposed models on mastery prediction tasks for using the dataset, which involved Math-Island online learning system, and 3) we explored the need to predict mastery in an online learning system autonomously. Our initial experiences have been very positive.

The present study has few limitations: This finding further emphasizes the necessity to use non-linear models to leverage historical data optimally. For solving time, the model has fit into the actual data very well; however, it seems more normalization tactics are needed for feedback request level. Furthermore, we used one learning unit's log data only; the result shows that more datasets and advanced deep learning and neural network models are needed, which is definitely included in our future research.

Although extending this study contains several possible directions to persevere. One such direction is incorporating these trained prediction models into a Math-Island tutoring system to estimate student mastery of skill and provide the amount of practice needed on time. Moreover, the interpretability of latent knowledge has not been fully explored, and further work is needed. For future work, we will not only explore how to design a mechanism, called, Mastery Prediction Model (MPM), to predict student mastery by adopting ML methods; but also investigate the following questions: 1) What is the minimum number of practice opportunities needed to achieve mastery of skill in mathematics learning; 2) Does the minimum opportunities vary by each domain in Math-Island; 3) After how many questions would it be suitable to predict students' mastery of skill, in order to make efficient adaptive learning context.

Acknowledgments

The authors would like to thank the Ministry of Science and Technology of the Republic of China, Taiwan, for financial support (MOST 109-2511-H-008 -011 -MY3), and Research Center for Science and Technology for Learning, National Central University, Taiwan. The authors would also like to thank all co-author for their assistance during this study.

References

Abidi, S. M. R., Hussain, M., Xu, Y., &Zhang, W. (2018). Prediction of confusion attempting algebra homework in an intelligent tutoring system through machine learning techniques for educational sustainable development. *Sustainability (Switzerland)*, *11*(1). https://doi.org/10.3390/su11010105

Ansems, E. L., Hanci, E., Ruijten, P. A. M., &IJsselsteijn, W. A. (2019). I focus on improvement: Effects of type of mastery feedback on motivational experiences. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *11433 LNCS*, 213–224. https://doi.org/10.1007/978-3-030-17287-9_18

Asselman, A., Khaldi, M., &Aammou, S. (2020). Evaluating the impact of prior required scaffolding items on the improvement of student performance prediction. *Education and Information Technologies*, *25*(4), 3227–3249. https://doi.org/10.1007/s10639-019-10077-3

Bälter, O., Zimmaro, D., &Thille, C. (2018). Estimating the minimum number of opportunities needed for all students to achieve predicted mastery. *Smart Learning Environments*, *5*(1). https://doi.org/10.1186/s40561-018-0064-z

Beck, J. E., &Gong, Y. (2013). Wheel-spinning: Students who fail to master a skill. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *7926 LNAI*, 431–440. https://doi.org/10.1007/978-3-642-39112-5-44

Caniëls, M. C. J., Chiocchio, F., &vanLoon, N. P. A. A. (2019). Collaboration in project teams: The role of mastery and performance climates. *International Journal of Project Management*, *37*(1), 1–13. https://doi.org/10.1016/j.ijproman.2018.09.006

Cen, H., Koedinger, K., &Junker, B. (2007). Is Over Practice Necessary ? – Improving Learning Efficiency with the Cognitive Tutor through Educational Data Mining. *Proceedings of the 13th International Conference on Artificial Intelligence in Education AIED 2007*, *158*, 511–518. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.142.7340&rep=%0Arep1&type=pdf

Chan, T.-W., Roschelle, J., Hsi, S., Kinshuk, Sharples, M., Brown, T., …Hoppe, U. (2006). One-To-One Technology-Enhanced Learning: an Opportunity for Global Research Collaboration. *Research and Practice in Technology Enhanced Learning*, *01*(01), 3–29. https://doi.org/10.1142/s1793206806000032

Corbett, A. T., &Anderson, J. R. (1994). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction*, *4*(4), 253–278. https://doi.org/10.1007/BF01099821

Gervet, T., Koedinger, K., Schneider, J., &Mitchell, T. (2020). When is Deep Learning the Best Approach to Knowledge Tracing? *Jedm*, *12*(3), 31–54.

Hussain, M., Zhu, W., Zhang, W., Abidi, S. M. R., &Ali, S. (2019). Using machine learning to predict student difficulties from learning session data. *Artificial Intelligence Review*, *52*(1), 381–407. https://doi.org/10.1007/s10462-018-9620-8

Imran, M., Latif, S., Mehmood, D., &Shah, M. S. (2019). Student academic performance prediction using supervised learning techniques. *International Journal of Emerging Technologies in Learning*, *14*(14), 92–104. https://doi.org/10.3991/ijet.v14i14.10310

Jin, C. (2020). MOOC student dropout prediction model based on learning behavior features and parameter optimization. *Interactive Learning Environments*, *0*(0), 1–19. https://doi.org/10.1080/10494820.2020.1802300

Lai, T. L., &Lin, H. F. (2015). A case study of the feedback design in a game-based learning for low achieving students. *Proceedings of the International Conference on E-Learning 2015, e-learning 2015 - Part of the Multi Conference on Computer Science and Information Systems 2015*, 213–214.

Lee, J. I., &Brunskill, E. (2012). The impact on individualizing student models on necessary practice opportunities. *Proceedings of the 5th International Conference on Educational Data Mining, EDM 2012*, 118–125.

Mao, Y., Lin, C., &Chi, M. (2018). Deep Learning vs. Bayesian Knowledge Tracing: Student Models for Interventions. *Journal of Educational Data Mining*, *10*(2), 28–54. Retrieved from https://jedm.educationaldatamining.org/index.php/JEDM/article/view/318

Myneni, L. S., Narayanan, N. H., Rebello, S., Rouinfar, A., &Pumtambekar, S. (2013). An interactive and intelligent learning system for physics education. *IEEE Transactions on Learning Technologies*, *6*(3), 228–239. https://doi.org/10.1109/TLT.2013.26

Ogdanova, D. B., Tova, Y. U. A. K., &Vetaeva, K. N. E. T. C. (2014). Experts’’ views consideration in assessing of the level of students’’ knowledge in distance learning. *Вестник Уфимского Государственного Авиационного Технического Университета*, *18*(5 (66)), 102–104.

Pavlik, P. I., Cen, H., &Koedinger, K. . (2009). Performance Factors Analysis – A New Alternative to Knowledge Tracing. *Materials Letters*, *212*, 531–538. Retrieved from http://dl.acm.org/citation.cfm?id=1659450.1659529

Pelánek, R. (2015). Metrics for Evaluation of Student Models. *JEDM - Journal of Educational Data Mining*, *7*(2), 1–19. Retrieved from http://www.educationaldatamining.org/JEDM/index.php/JEDM/article/view/JEDM087

Piech, C., Bassen, J., Huang, J., Ganguli, S., Sahami, M., Guibas, L., &Sohl-Dickstein, J. (2015). Deep knowledge tracing. *Advances in Neural Information Processing Systems*, *2015-Janua*, 505–513.

Razzaq, R., Ostrow, K. S., &Heffernan, N. T. (2020). Effect of Immediate Feedback on Math Achievement at the High School Level. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/978-3-030-52240-7_48

Romero, M., Hernández, J. M., Juola, J. F., Casadevante, C., &Santacreu, J. (2020). Goal Orientation Test: An Objective Behavioral Test. *Psychological Reports*, *123*(4), 1425–1451. https://doi.org/10.1177/0033294119845847

Rong, S., &Bao-Wen, Z. (2018). The research of regression model in machine learning field. *MATEC Web of Conferences*, *176*, 8–11. https://doi.org/10.1051/matecconf/201817601033

Schnipke, D. L. (2002). Exploring Issues of Examinee Behavior: Insights Gained from Response-Time Analyses. *Computer-Based Testing: Building the Foundation for Future Assessments*, 237–266.

Slater, S., &Baker, R. (2019). Forecasting future student mastery. *Distance Education*, *40*(3), 380–394. https://doi.org/10.1080/01587919.2019.1632169

Sokkhey, P., &Okazaki, T. (2019). Comparative Study of Prediction Models for High School Student Performance in Mathematics. *IEIE Transactions on Smart Processing & Computing*, *8*(5), 394–404. https://doi.org/10.5573/ieiespc.2019.8.5.394

Sokkhey, P., &Okazaki, T. (2020). Hybrid machine learning algorithms for predicting academic performance. *International Journal of Advanced Computer Science and Applications*, *11*(1), 32–41. https://doi.org/10.14569/ijacsa.2020.0110104

Tang, H., Xing, W., &Pei, B. (2018). Exploring the temporal dimension of forum participation in MOOCs. *Distance Education*, *00*(00), 1–20. https://doi.org/10.1080/01587919.2018.1476841

Tuominen, H., Juntunen, H., &Niemivirta, M. (2020). Striving for Success but at What Cost? Subject-Specific Achievement Goal Orientation Profiles, Perceived Cost, and Academic Well-Being. *Frontiers in Psychology*, *11*(September), 1–18. https://doi.org/10.3389/fpsyg.2020.557445

Yeh, C. Y. C., Cheng, H. N. H., Chen, Z. H., Liao, C. C. Y., &Chan, T. W. (2019). Enhancing achievement and interest in mathematics learning through Math-Island. *Research and Practice in Technology Enhanced Learning*, *14*(1). https://doi.org/10.1186/s41039-019-0100-9

Yu, L. C., Lee, C. W., Pan, H. I., Chou, C. Y., Chao, P. Y., Chen, Z. H., …Lai, K. R. (2018). Improving early prediction of academic failure using sentiment analysis on self-evaluated comments. *Journal of Computer Assisted Learning*, *34*(4), 358–365. https://doi.org/10.1111/jcal.12247