A MOOC-Based Hybrid Teaching Model of College English

https://doi.org/10.3991/ijet.v18i02.35535

Tan Feng[1(✉)], Laith Abualigah[2,3,4]
[1]School of Humanities, Geely University of China, Jianyang, China
[2]Faculty of Computer Sciences and Informatics, Amman Arab University, Amman, Jordan
[3]Faculty of Information Technology, Middle East University, Amman, Jordan
[4]School of Computer Sciences, Universiti Sains Malaysia, Pulau Pinang, Malaysia
18407274@masu.edu.cn

Abstract—In the era of intelligence, Internet + technology is widely used in various fields, and English Teaching in the education industry of colleges and universities gradually tends to be an online and offline mixed teaching mode. However, under the MOOC model, the feedback of College Students' English learning and the recognition of their knowledge level has become new difficulties. Aiming at the feedback of students' learning situation under the mixed mode of College English teaching, this paper uses the optimized Bayesian knowledge tracking model (BKTM) to predict students' English learning situation and introduces students' learning behavior and forgetting behavior to optimize parameters. Finally, a performance verification experiment is carried out by analyzing the students' answer performance in College English mixed teaching. The results show that the prediction errors of the four knowledge points of 60 students in the two classes are all about 7%, and the maximum error is 11%. Experiments show that the model has high accuracy and stable performance in predicting the probability of mastering knowledge points.

Keywords—English class, mixed teaching mode, Bayesian network, Knowledge tracking model

1 Introduction

In the mixed English teaching model in colleges and universities, understanding students' teaching evaluation and learning is an indispensable part of the course [1]. However, in MOOC English Teaching in Colleges and universities, students and teachers cannot interact in time, resulting in the problems of backward teaching progress and low learning efficiency [2]. Because of the difficulties of students' learning feedback in the MOOC-based online and offline hybrid teaching mode of College English, this paper studies the construction of a prediction model of students' learning status in English teaching based on the Bayesian network [3]. In order to improve the prediction accuracy of the Bayesian knowledge tracking model, the learning behavior and forgetting behavior of students in MOOC are introduced. The behavior parameters are quantified by the decision tree algorithm, the maximum expectation algorithm obtains the optimal probability parameters. The innovation of this study is that after optimizing the parameters of learning behavior and forgetting behavior, the Bayesian network can make personalized predictions and suggestions for students' English knowledge point structure, helping improve students' learning methods and teaching quality.

2 Journals reviewed

At the beginning of the Bayesian network, its models were mainly applied to the probability reasoning of artificial intelligence. Since the 1990s, different scholars have gradually improved various algorithms of Bayesian networks. Regarding data mining and knowledge statistics, Duan et al. Proposed a heuristic algorithm based on mutual information to fuse tree augmented naive Bayesian network to build a model, which is used to dynamically describe the dependencies in unlabeled test cases in data mining. Experimental results show that the model is superior to K-means Bayesian classifier and random forest algorithm [4].

Chen et al. Used Bayesian networks to build a probabilistic decision-making model for the rear-end collision avoidance system. Experiments show that the model considers the impact of vehicle dynamics, driver reaction ability, and external environment on rear-end collision and can accurately estimate the rear-end collision risk with relatively low complexity [5]. Wang l proposed a Bayesian network classifier learning method that fully represents the diversity of conditional dependencies. The results show that the classification performance of the algorithm is still competitive compared with the most advanced single model learner and integrated learner [6]. Zhang et al. Used a Bayesian network to analyze the efficacy and safety of excellent saphenous vein transplantation (BSV) bypass surgery in treating femoral-popliteal artery occlusive diseases. The all-cause death data within five years proved that the main complications of BSV were not related to death [7]. Eschweiler et al. Also used a Bayesian network meta-analysis of randomized controlled trials to analyze the clinical effects of non-steroidal anti-inflammatory drugs and gabapentin drugs on chronic low back pain [8], Soltanali et al. Used Bayesian network (BN) theory to determine the nondynamic relationship between complex system failures in FFTA model. The results show that the BN model considers conditional rules to reflect the dynamic relationship between faults [9].

Zhao et al. Used a dynamic Bayesian network to predict visibility in the Arctic based on the physical basic statistical model. Experiments have proved that using a dynamic Bayesian network to predict visibility is adequate for predicting climate change in the Arctic [10]. Zhou et al. Extended Bayesian network feature Finder (Banff), which provides a graphical description package of posterior reasoning, model comparison, and model fitting for understanding biological mechanisms and disease pathology [11].

Geng et al. Used a Bayesian algorithm and composite importance measure to rank the factors affecting survival and built a prediction model for advanced gallbladder cancer (GBC). Experiments show that the BN Based survival prediction model can be used as a decision support tool for patients with advanced GBC [12]. Boya et al. Used Bayesian network meta-analysis to prove the efficacy and safety of opioid analgesics in the treatment of chronic low back pain. The results showed that the effect of

Bayesian NMA was estimated in the form of odds ratio (or) and 95% confidence interval (CRI) [13]. In order to improve the accuracy of fault cause diagnosis in chemical processes, Kumari et al. Proposed an improved Bayesian network (MBN), which considered the cycle. At the same time, the Tennessee Eastman process is taken as an example to verify the method's effectiveness. [14]. Xue et al. Transformed the belief rule base into the Bayesian network's conditional probability table (CPT) to express the influence relationship of various factors in the location selection of offshore wind turbines. The model's effectiveness is verified by taking the busy waterway in the East China Sea as an example [15].

Qi et al. Proposed an LVS-based BN parameter dynamic embedded learning method to represent unobservable abstract concepts such as patient disease and customer credit. The results show this method's efficiency, convergence, and accuracy are better than other LVS BN parameter learning methods [16]. Noriaki et al. Developed an R software package CBNPLOT to infer Bayesian networks (BNs) from gene expression data, explicitly utilizing the functional enrichment analysis (EA) results obtained from the managed biological pathway database. Experiments show that this method is absolutely effective in evaluating and analyzing established knowledge and literature, and it is possible to promote the discovery of knowledge from gene expression data sets. [17].

Nguyen et al. Developed an agricultural robot system (ARS/PSPD) for plant stress propagation detection based on the Bayesian network scanning protocol. Experiments show that Bayesian network inference's scanning protocol is superior to all other protocols. [18]. Chen Z proposed an unsupervised Bayesian network ensemble (BNI) method to detect driving genes and estimate the spread of disease at the patient and/or cohort levels based on accurately located differentially expressed genes, variants in somatic mutations, and gene interaction networks. Experiments show that this method can capture the inherent structural information [19].

To sum up, the current education is constantly in line with computer technology, but the application results of technology are still insufficient, and students' learning status and results cannot be timely fed back. The Bayesian network performs better in probability reasoning of uncertain events, so it is widely used in clinical medicine, plant protection, software development, and other fields. At this stage, the technical application of online teaching mode is not mature enough, and the students' learning status tracking and knowledge learning level are not well understood. In the research of using its precise probability prediction to establish the student status prediction of College English mixed mode, it is expected to help college English teaching establish more effective teaching methods and improve students' learning ability and enthusiasm.

3 Construction of bayesian knowledge tracking model for mixed english teaching

3.1 Bayesian model analysis based on learning prediction algorithm

A bayesian network is constructed based on the Bayesian formula and can be used to deal with probability events in uncertain information. The student learning state prediction mechanism of the Bayesian knowledge tracking model based on the hybrid teaching model is based on the calculation of state transition probability. After observing and calculating the probability distribution of the knowledge node to the following knowledge node and the corresponding state node, the model can predict the learning situation of College Students' English knowledge. Therefore, in essence, the Bayesian knowledge tracking model is a unique hidden Markov model (HMM) [20]. Therefore, the principle of applying the Bayesian network to predict students' learning status under the College English mixed teaching mode course is shown in Figure 1.

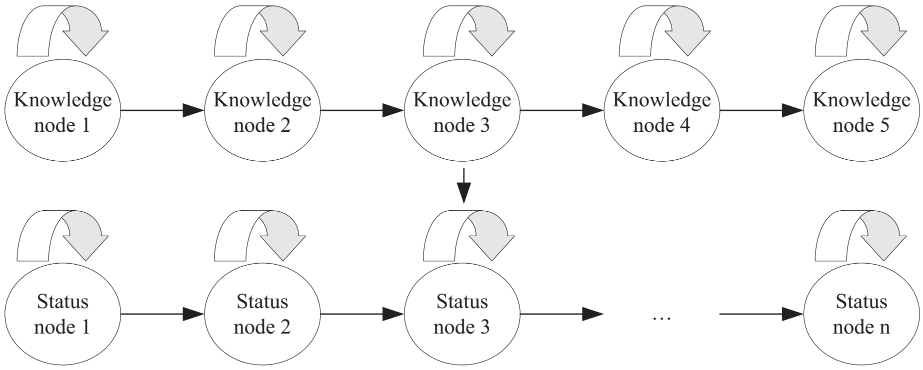

Fig. 1. The principle of applying the Bayesian network to students' learning state prediction

The changes in knowledge nodes and student status nodes in the Bayesian knowledge tracking model in Figure 1 are introduced by students' learning behavior. The probability model in the figure is based on the hidden Markov time series model, which describes the randomly generated state sequence and generates the observable sequence from the random state sequence [21]. According to the Markov chain structure in the figure, the following probability formula (1) is obtained.

$$P(S_{n+1} = s \,|\, S_0, S_1, S_2, ..., S_n) = P(S_{n+1} = s \,|\, S_n) \tag{1}$$

In formula (1), P is the calculated state transition probability, S is the student learning state node, and n is the state node number. When applied to students' learning state detection, all knowledge nodes in BKTM are constructed based on the hierarchical relationship of College English teaching content, and the probability of knowledge points' mastery state in the model is judged according to the correct answer rate. However, as students' learning behavior in the mixed teaching mode of College English is the main factor affecting the state transition probability, the student learning characteristic data is extracted to introduce behavior nodes into the model. In BKTM, the data of knowledge nodes can be quantified as "master knowledge" and "master knowledge", and the quantification of behavior data of performance nodes can also be set as "correct answer" and "wrong answer". At the same time, the quantification of behavior data depends on the decision tree algorithm. After preprocessing the data, the quantitative, mixed mode learning behavior can improve the answer performance by 1, while 0 means that the learning behavior hurts the answer performance [22]. After introducing learning behavior, the probability calculation formula of correct answer performance is shown in formula (2).

$$P(r) = P(K_n)((1 - P(W)) + (1 - P(R)))P(R) + P(R)(1 - P(W)) \quad (2)$$

In equation (2), $P(r)$ represents the probability of correct answer, $P(K_n)$ represents the probability of knowledge node n mastering, $P(W)$ represents the correct rate of guessing answer behavior without knowledge points, and $P(R)$ represents the error rate of answer behavior after mastering knowledge points. After the introduction of learning behavior, the correct performance of answering questions can be divided into four situations: whether the students have mastered the knowledge points, whether the learning behavior has a positive impact, and whether the wrong behavior has a negative impact. Similarly, the Bayesian formula for the probability of wrong answers after introducing learning behavior is shown in formula (3).

$$P(w) = P(K_n)P(W) + (1 - P(K_n))(1 - P(R)) + (1 - P(R))P(W) \quad (3)$$

In equation (3), $P(w)$ is the probability of a wrong answer. The correct and wrong sequence of answer results is obtained according to the correct and wrong answer probability in the behavior BKTM to calculate the learning prior probability of the Bayesian network. The calculation formula is shown in equation (4).

$$\begin{cases} P(K|r) = \dfrac{P(K)(1 - P(W)) + P(K)P(R)(1 - P(W))}{P(K)(1 - P(W)) + (1 - P(K))P(R) + P(R)(1 - P(W))} \\ P(K|w) = \dfrac{P(K)(1 - P(R))(PW) + P(K)P(R)}{P(K_n)P(W) + (1 - P(K_n))(1 - P(R)) + (1 - P(R))P(W)} \end{cases} \quad (4)$$

In equation (4), $P(K|r)$ represents the prior probability of students' knowledge point learning when the answer is correct and $P(K|w)$ represents the prior probability of students' knowledge point learning when the answer is wrong. According to the performance of the two answer States, update the mastery probability of knowledge points according to the formula $P(K) = P(K|r) + P(K|w)$. Finally, the BKTM with learning behavior is introduced to predict the mastery probability of the next knowledge node according to the updated mastery probability, and its calculation formula is shown in formula (5).

$$P(K_{n+1}) = P(K_n) + (1 - P(K_n))P(T) \quad (5)$$

In equation (5), $P(K_n)$ represents the mastery probability of the knowledge node n, $P(K_{n+1})$ represents the mastery probability of the next knowledge node $n + 1$, and $P(T)$ represents the learning transfer probability of students.

3.2 Parameter optimization of Bayesian knowledge tracking model

Under the online and offline mixed teaching mode of College English, learning behavior is the main influencing factor of students' learning. However, college students' forgetting behavior in the learning process cannot be ignored. In the Bayesian knowledge tracking model, students' performance nodes are also affected by forgetting.

In reality, knowledge points need to be memorized repeatedly to avoid forgetting. Therefore, to increase the prediction accuracy of BKTM (Bayesian Knowledge Tracking Model), students' forgetting behavior needs to be quantified as forgetting probability and introduced into the model. The learning and forgetting behaviors are introduced into the structural parameters of BKTM simultaneously, as shown in Figure 2.

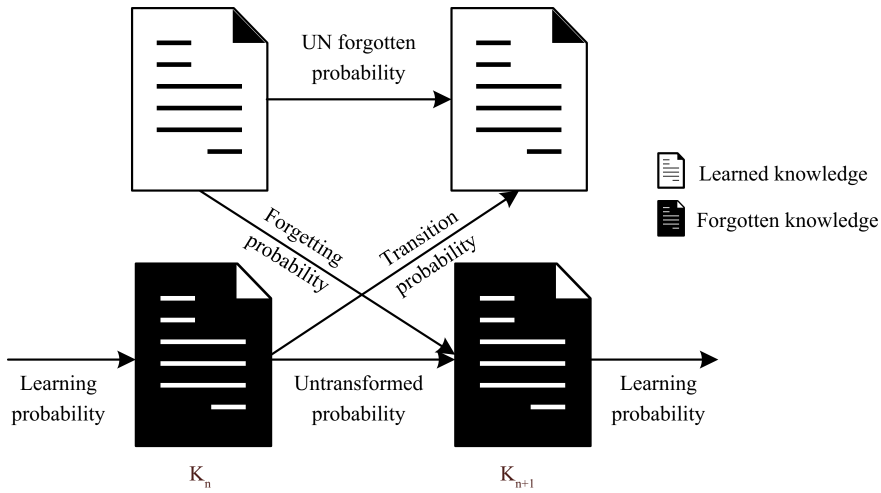

Fig. 2. Learning behavior and forgetting behavior are introduced into the structural parameters of BKTM

It can be seen from Figure 2 that after the Bayesian knowledge tracking model is introduced into the students' forgetting behavior and learning behavior at the same time. The correct or wrong answer performance nodes will be determined by the error probability and guess probability in the learning behavior parameters and the forgetting probability in the forgetting behavior parameters simultaneously. After introducing the forgetting parameter, the correct performance of students' English knowledge points can be divided into four types: knowledge points mastery, negative learning behavior, and low probability of error and forgetting; Having mastered knowledge points, active learning behavior, and low probability of error forgetting; They do not master knowledge, have positive learning behavior, and have a certain probability of forgetting; They did not master the knowledge points, and their learning behavior was negative, but their error rate was low, and their guesses were correct. Therefore, after the forgetting probability in the Bayesian model, the probability calculation formula of the correct performance of students' answers is shown in formula (6).

$$P(r) = P(K_n)(1-P(W))(1-P(F)) + P(R)(1-P(F)) + (1-P(K_n))P(R) \tag{6}$$

In equation (6), $P(F)$ represents the forgetting probability of students' knowledge points. In the same way as the correct answer, under the joint action of learning behavior and forgetting behavior, there are four types of the wrong answer, namely, having mastered knowledge, active learning behavior, and high probability of forgetting, and mistakes; Having mastered the knowledge points, the learning behavior has a negative impact, resulting in a high probability of forgetting; They do not master the knowledge and have positive learning behavior, but they have the transformation of forgetting and positive influence of behavior; Lack of knowledge, negative learning behavior and low accuracy of guessing. Bring the above four situations into the Bayesian formula to obtain the calculation formula of students' answer error probability, and its mathematical expression is shown in formula (7).

$$P(w) = P(K_n)(1-P(W))(1-P(F)) + (1-P(R)P(W)P(F) + (1-P(K_n))(1-P(R)) \quad (7)$$

According to the correct and wrong answer sequence of the Bayesian knowledge tracking model with forgotten parameters, the learning prior probability of the model is optimized, and the formula is shown in formula (8).

$$\begin{cases} P(K|r) = \dfrac{P(K)P(R)(1-P(F)) + P(K)(1-P(W))(1-P(F))}{P(K_n)(1-P(W))(1-P(F)) + P(R)(1-P(F)) + (1-P(K_n))P(R)} \\ P(K|w) = \dfrac{P(K)P(W)P(F)(1-P(R)) + P(K)P(W)P(F)}{P(K_n)(1-P(W))(1-P(F)) + (1-P(R)P(W)P(F)} \\ \qquad\qquad + (1-P(K_n))(1-P(R)) \end{cases} \quad (8)$$

Finally, after completing all the answers according to BKTM and optimizing the learning probability, the formula for predicting the probability of students' mastering the next knowledge point is shown in formula (9) according to the mastery of knowledge points and the performance of the answers.

$$P(K_{n+1}) = P(K_n)(1-P(F)) + (1-P(K_n))P(T)(1-P(F)) \quad (9)$$

After introducing the learning behavior parameters and forgetting behavior parameters of students, the prediction accuracy of the Bayesian knowledge tracking model can be improved in theory. For the evaluation of the prediction effect of the model, the root means square error (RMSE) is used to express the numerical dispersion of the model, and its calculation formula is shown in equation (10).

$$R = \sqrt{\frac{\sum_{i=1}^{n}(P_i - p_i)^2}{n}} \quad (10)$$

In equation (10), R represents the root mean square error of the model, i represents the prediction times, n represents the maximum number of prediction samples, P_i represents the model prediction value, p_i represents the actual probability value, and $P_i - p_i$ represents the deviation between the model prediction probability and the actual probability value [23]. In the Bayesian knowledge tracking model, the root means square error value represents the ratio of the square of the error between the prediction result of the model and the students' learning and mastering situation to the prediction times [24]. When the calculation result of root mean square error is small, it indicates that the prediction result of the model has high accuracy. When the calculation result of root mean square error is small, it indicates that the prediction effect of the model on students' learning is poor, and the error is significant. After the optimized model is obtained, the process of applying the BKTM model and predicting students' situation in the mixed teaching mode of College English is summarized as shown in Figure 3.

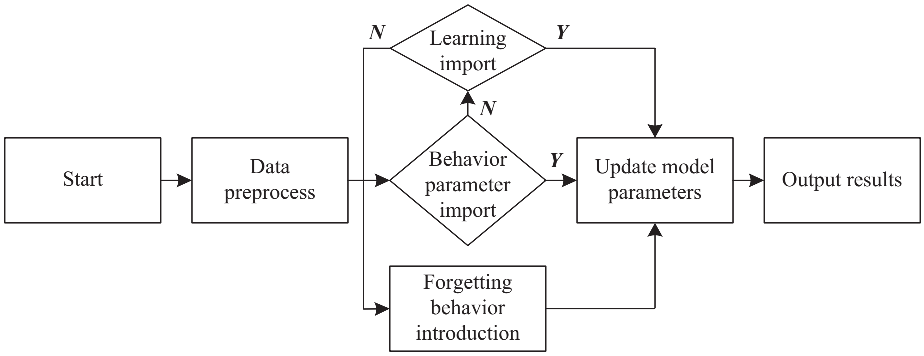

Fig. 3. The application of the BKT model and the process of student situation prediction of college English mixed teaching model

It can be seen from Figure 3 that the data preprocessing methods are divided into simple quantization and decision tree quantization. For knowledge nodes and performance nodes, 0 and 1 of simple quantization are used to represent errors and correctness, while for learning behavior, the decision tree algorithm is used to quantify the positive impact of learning behavior as one and the negative impact as 0. After introducing learning parameters and forgetting parameters, the model's correct and wrong answer sequences are obtained, and the Bayesian formula is updated. Students' correct and wrong answer sequences are used for model training, and the best parameters are obtained through the maximum expectation algorithm. Finally, the model's prediction results are evaluated using the root mean square error [25].

4 Performance analysis of Bayesian knowledge tracing model applied to students' learning state prediction

4.1 Simulation training and optimal parameter solution of Bayesian knowledge tracking model

In order to verify the accuracy of the Bayesian knowledge tracking model after introducing students' learning behavior parameters and students' forgetting behavior parameters in the prediction of students' learning state, the available data sets bridge to algebra and school data with the effect provided by the online teaching platform of assistants are used to simulate and train the model. The data processing and picture drawing tools used in this experiment are EXCEL table and VISIO2016 software respectively. The basic parameters of the two public datasets are shown in Table 1.

Table 1. Expose the parameters of the bridge to algebra and school data with effect

Data Set	Total Data (Thousand)	Number of Valid Samples (Thousand)	Number of Initial Data Items	Number of Valid Data Items
Bridge to Algebra	813.6	321.0	19	7
School Data with Affect	1048.6	449.9	35	8

In the above two data sets, the valid data items in bridge to algebra include seven kinds of valid data items: line number, student D, knowledge chapter, question name, correct or wrong answer, and knowledge investigated by the question, while the school data with effect data set includes eight kinds of valid data items: answer time, request prompt, student ID, knowledge chapter, question name, correct or wrong answer, and knowledge investigated by the question. As the classification of the bridge to algebra's professional knowledge is not as straightforward as that of school data with effect, the best parameters of the model are first calculated by using the school data with effect data set training and the maximum expectation algorithm. MATLAB completes the simulation training, and the maximum number of training is 600. Then the likelihood changes of the five probability parameters are shown in Figure 4.

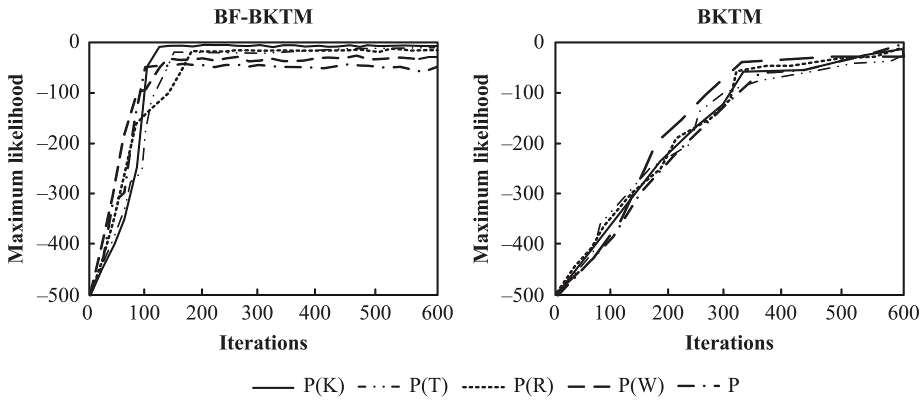

Fig. 4. Likelihood variation of probability parameters in BKTM and BF-BKTM algorithms

As seen from Figure 4, when the number of iterations of the BF-BKTM model with learning behavior and forgetting behavior is close to 120, the likelihood probability tends to be flat, indicating that the maximum likelihood probability has been reached. In contrast, when the number of iterations of the non-optimized BKTM model is close to 40, the likelihood probability tends to be flat, indicating that the maximum likelihood probability has been reached. Compared with the iterative training effect of the traditional BKTM model and the BF-BKTM model in the data set school data with effect, it is evident that the optimized model has higher accuracy in the prediction of answer performance. The best parameters of the model calculated according to the training optimization model and the maximum expectation algorithm are shown in Table 2.

Table 2. Optimal probability parameters of BF-BKTM

Probability Parameter	P(K)	P(T)	P(R)	P(W)	P(F)
Initial value	0–0.999	0–0.999	0–0.500	0–0.500	0–0.999
Optimum parameters	0.594	0.327	0.266	0.317	0.616

In Table 2, P(K) represents the probability of mastering knowledge nodes, P(T) represents the learning transfer probability of students, P(W) represents the correct rate of guessing and answering behavior without mastering knowledge points, P(R) represents the wrong rate of answering behavior after mastering knowledge points, and P(F) represents the probability of forgetting knowledge points of students. The purpose of setting the initial values of guess probability and error probability to 0–0.5 is to reduce the degradation of the model. In order to verify the validity of the correlation application between the answer performance samples in the data set and the model, the data set school data with effect is processed in segments, and 50 points are randomly selected from the data samples that have not entered the training simulation for model prediction. The ratio of the actual probability to the prediction probability of the model is taken as the prediction accuracy. If the ratio is close to 1 in a month, the higher the model's prediction accuracy is, and the prediction accuracy of the three algorithm models is shown in Figure 5.

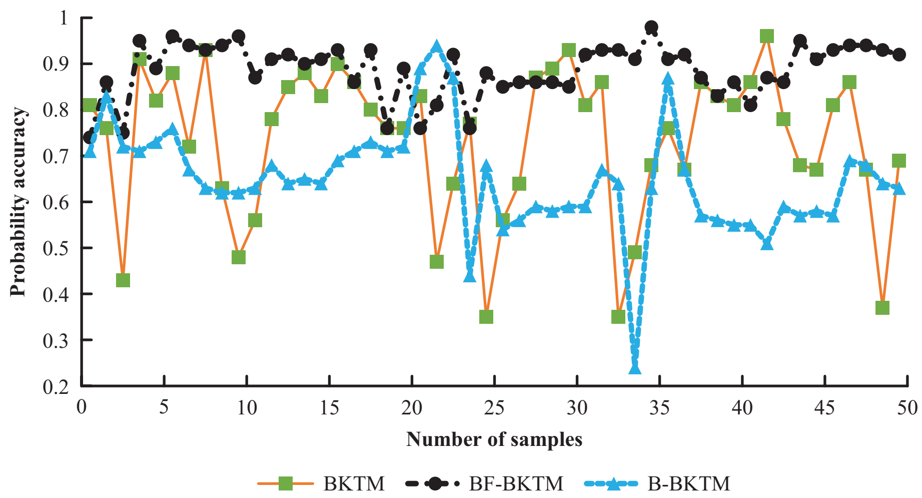

Fig. 5. Prediction accuracy of student state based on different algorithms

In Figure 5, BKTM represents the initial model, B-BKTM represents the model after introducing students' learning behavior, and BF-BKTM represents the Bayesian knowledge tracking model with students' learning and forgetting behavior. As can be seen from Figure 5, BF-BKTM has the most stable and accurate prediction results among the three algorithms. Among the prediction results of 50 test samples, the average prediction accuracy of the three models from high to low is 88.7%, 73.5%, and

64.9%, corresponding to BF-BKTM, B-BKTM, and BKTM, respectively. Experiments show that BF-BKTM, which introduces learning and forgetting behavior, has the highest accuracy in predicting students' answer performance. Similarly, the bridge to algebra dataset is processed in sections, 50 answer samples are randomly selected, and the test samples of the school data with effect dataset remain unchanged. Finally, the root mean square error of the three algorithm models is calculated, and the results are shown in Figure 6.

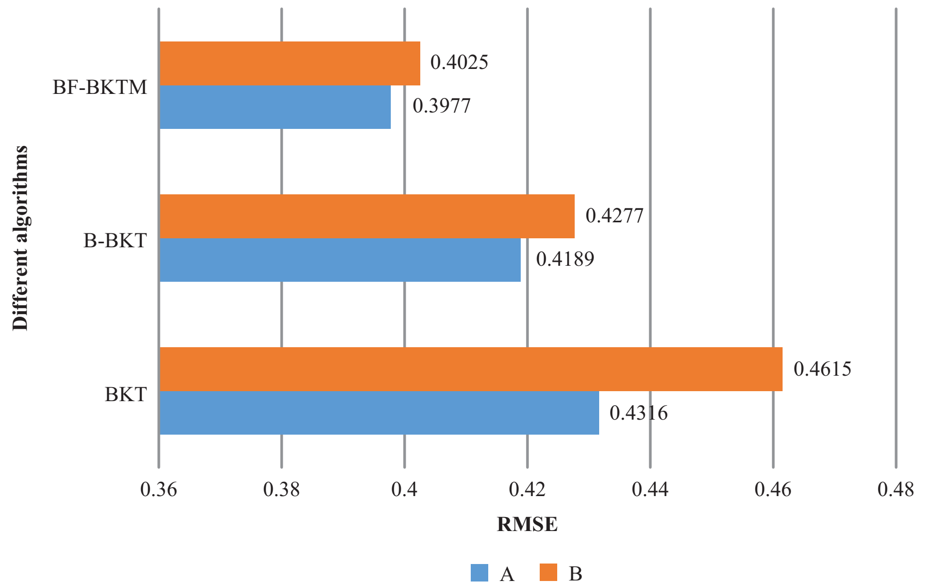

Fig. 6. Prediction root mean square error of different algorithms in two data sets

Figure 6, a represents the school data with the effect dataset, and B represents the bridge to algebra dataset. As seen from Figure 6, the root means the square error of the three algorithm models in the school data with effect dataset is smaller than that in the bridge to algebra dataset. At the same time, the root means a square error of BF-BKTM with learning behavior, and forgetting behavior is the lowest in the three models. In dataset a, the RMSE of BF-BKTM is 0.3977, while in dataset B, the RMSE is 0.4025. Simulation results show that BF-BKTM with learning and forgetting behavior has the highest accuracy in predicting students' learning state.

4.2 Performance analysis of Bayesian knowledge tracing model applied to mixed mode courses in Colleges and Universities

The subjects of this study were 61 students in four homogeneous classes of sophomores in an undergraduate university, including 11 boys and 17 girls in class A, a total of 28; In class B, there are 12 boys and 20 girls of 32. By investigating the College English Teaching Indicators and the questions involved in the English proficiency test, the study designed 50 questions covering four English abilities: listening comprehension, reading comprehension, grammar, and translation. The test questions are multiple-choice questions and blank filling questions. The correct and wrong answers are simple quantized bits 1 and 0. The parameters of students' learning status test papers are shown in Table 3.

Table 3. Content and score of test paper

Test Items	Question Type	Question Quantity	Fraction
Listening comprehension	Multiple choice	15	30
Reading comprehension	Multiple choice/Completion	15	30
Grammar	Multiple choice/Completion	10	20
Translate	Completion	10	20

Number the four knowledge types in the table, set listening comprehension as knowledge K1, reading comprehension as knowledge K2, grammar as knowledge K3, and translation as knowledge K4. The answer time of the test paper is 60 minutes. The test scores of the two classes are shown in Figure 7.

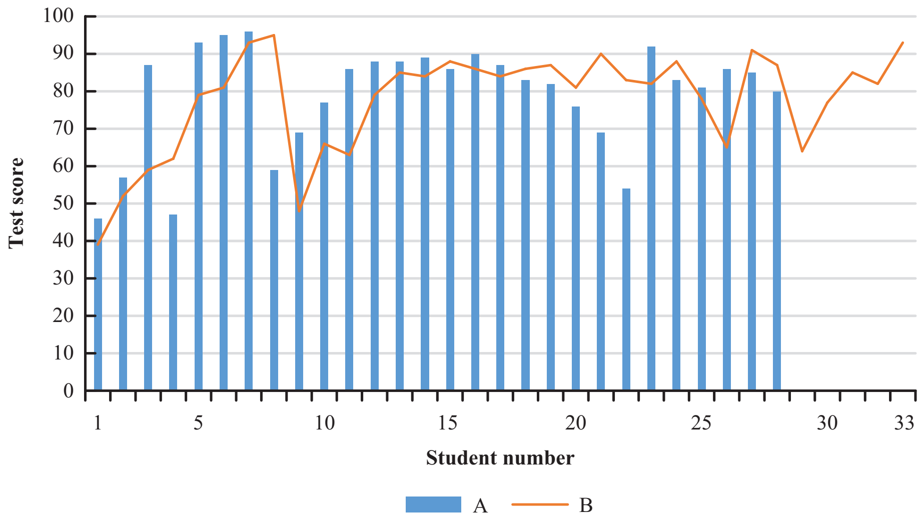

Fig. 7. Test scores for different classes

As seen in Figure 7, most students' scores in classes A and B are distributed in the range of 70–95. A. B. In the two classes, the number of students with a score below 60 is 5 and 4, respectively. The number of people with a 60–80 is 5 and 8, respectively, and those with a score higher than 80 are 18 and 20, respectively. The Bayesian knowledge tracking model constructed by the research is used to analyze the students' answer data, and the knowledge mastery of 60 students is obtained. If the probability of mastering the obtained knowledge points is less than 0.5, it will be regarded as not mastering; 0.5–0.9 means a good mastery, and more than 0.9 means an excellent mastery. The overall 60 students' mastery of the four knowledge points is shown in Figure 8.

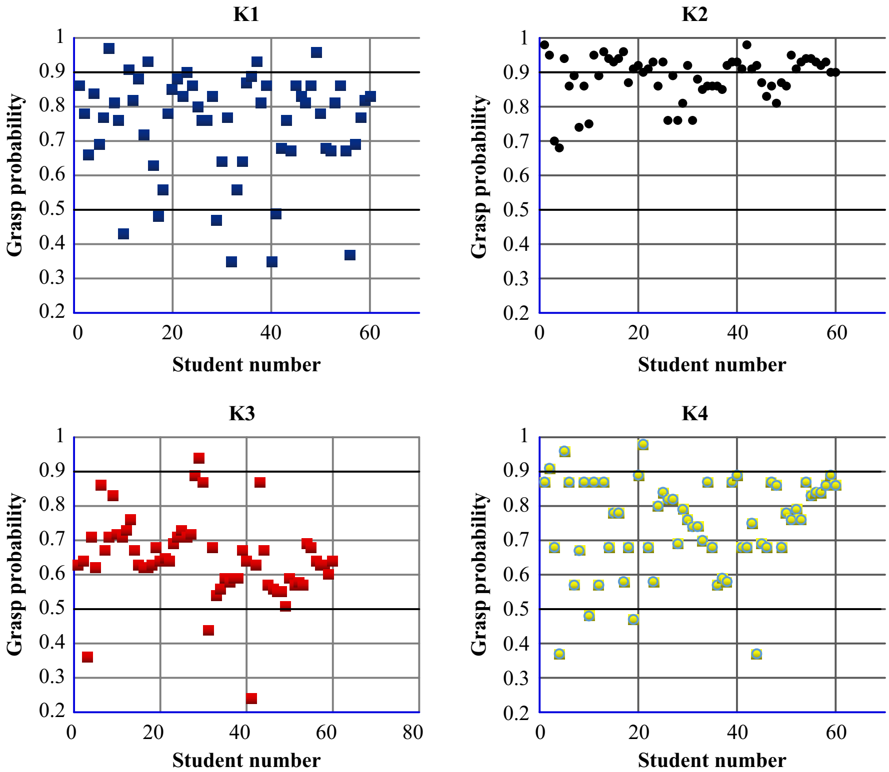

Fig. 8. The mastery of the four knowledge points by all students in the two classes

It can be seen from Figure 8 that all 60 students in the two classes have the highest mastery of the four knowledge points in the reading test. Through the Bayesian knowledge tracking model, students' average probability of mastering knowledge points is 0.76 for listening comprehension (K1), 0.88 for reading comprehension (K2), 0.68 for grammar knowledge points (K3), and 0.77 for translation knowledge points (K4). There are seven students whose listening comprehension (K1) Mastery probability is below 0.5, 48 students with a good mastery degree, and five whose listening comprehension (K1) Mastery probability is above 0.9. There are 0 students whose reading comprehension (K2) Mastery probability is below 0.5, 30 students whose mastery probability is within the range of 0.5–0.9, and the other 30 students' mastery probability is above 0.9. There are three students whose probability of mastering grammar knowledge points (K3) is less than 0.5, 56 students who master grammar well, and one student whose probability of mastering grammar knowledge points (K3) is more than 0.9. There are four students whose mastery probability of translation knowledge points (K4) is below 0.5, 53 students whose mastery probability is within the range of 0.5–0.9, and the other three students' mastery probability is above 0.9. Finally, by comparing the probability of knowledge points predicted by the model with the probability of actual answer performance, the accuracy of the Bayesian knowledge tracking model with learning behavior and forgetting behavior is verified. The comparison between the predicted and actual results of the two classes is shown in Figure 9.

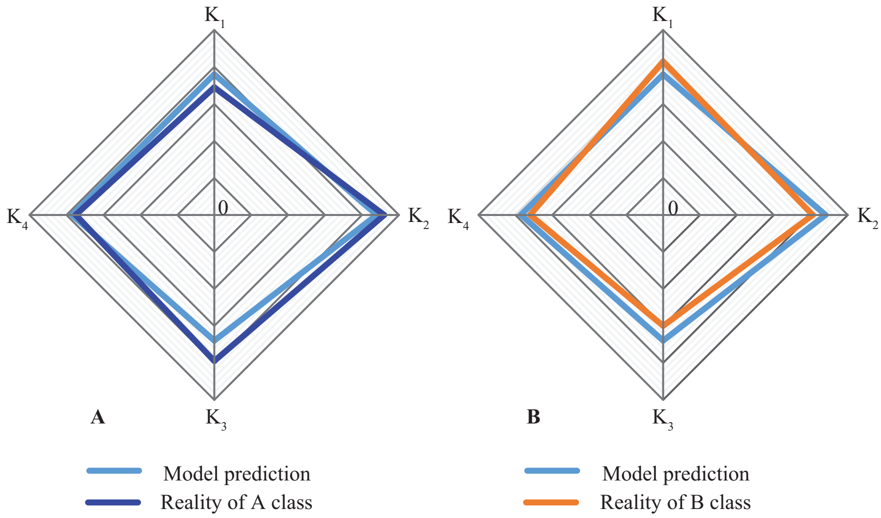

Fig. 9. Comparison between the predicted results of the two classes and the actual situation

It can be seen from Figure 9 that there is a small gap between the model prediction and the actual probability of mastering the four types of knowledge points of students in classes A and B. In class A, the error of listening comprehension mastery probability is 7%, and the error of reading comprehension mastery probability is 4%. In class A, the probability error of mastering grammar knowledge points is the largest, with a value of 11%; However, the error of mastering translation knowledge points is the smallest, which is 2%. In the prediction of class B, the probability error of listening comprehension and reading comprehension is 7%. In comparison, the maximum error is 8% in the probability of mastering grammar knowledge points, and the minimum error is 5% in the probability of mastering translation knowledge points. The overall error comparison data shows that the error of the four knowledge points of 60 students in the two classes is about 7%, and the maximum error is 11%. The experimental results show that the accuracy performance of the knowledge point mastery probability predicted by the model is high.

5 Conclusion

In order to solve the problem of teachers' low mastery of students' learning in the mixed teaching mode of College English, this paper studies the optimization of BKTM by using students' learning behavior and forgetting behavior and carries out the simulation training and practical application of the model. In the sample test of school data with effect dataset, the average prediction accuracy of BF-BKTM, B-BKTM, and BKTM is 88.7%, 73.5%, and 64.9% from high to low; The root mean square errors were 0.3977, 0.4189 and 0.4316 respectively. In the practical application of the two experimental classes, the average probabilities of students' mastery of knowledge points in the Bayesian knowledge tracking model are listening comprehension (K1)

0.76, reading comprehension (K2) 0.88, grammar knowledge points (K3) 0.68, and translation knowledge points (K4) 0.77; The error between the predicted probability and the actual probability of the two classes is about 7%, and the maximum error is 11%. Experiments show that BF-BKTM, which introduces learning and forgetting behavior, has high accuracy and robustness in predicting students' learning status and answer performance. The deficiency of this experiment is that the total number of students applying the experiment is 60, and the sample situation is not enough to replace the learning situation of all students.

6 References

[1] M. Aparicio, T. Oliveira, F. Bacao, et al. "Gamification: A Key Determinant of Massive Open Online Course (MOOC) Success", *Information & Management*, vol. 56, pp. 39–54. 2019. https://doi.org/10.1016/j.im.2018.06.003

[2] G. Nanda, K. A. Douglas, D. R. Waller, et al. "Analyzing Large Collections of Open-Ended Feedback from MOOC Learners Using LDA Topic Modeling and Qualitative Analysis", *IEEE Transactions on Learning Technologies*, vol. 2, pp. 146–160, 2021. https://doi.org/10.1109/TLT.2021.3064798

[3] P. Librizzi, A. Biswas, R. Chang, et al. "Broadband Chiral Hybrid Plasmon Modes on Nanofingernail Substrates", *Nanoscale*, vol. 12, pp. 3827–3833, 2020. https://doi.org/10.1039/C9NR07394A

[4] Z. Duan, L. Wang, M. Sun, "Efficient Heuristics for Learning Bayesian Network from Labeled and Unlabeled Data", *Intelligent Data Analysis*, vol. 24, pp. 385–408, 2020. https://doi.org/10.3233/IDA-194509

[5] X. M. Chen, et al. "A Rear-End Collision Risk Evaluation and Control Scheme Using a Bayesian Network Model", *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, pp. 264–284, 2019. https://doi.org/10.1109/TITS.2018.2813364

[6] L. Wang, P. Chen, S. Chen, et al. "A Novel Approach to Fully Representing the Diversity in Conditional Dependencies for Learning Bayesian Network Classifier", *Intelligent Data Analysis*, vol. 25, pp. 35–55, 2021. https://doi.org/10.3233/IDA-194959

[7] R. Zhang, L. Ni, R. Zeng, et al. "An Indirect Comparison by Bayesian Network Meta-analysis of Drug-coated Devices versus Saphenous Vein Graft Bypass in Femoropopliteal Arterial Occlusive Disease", *Journal of Vascular Surgery*, vol. 2, pp. 478–486, 2021. https://doi.org/10.1016/j.jvs.2020.11.054

[8] E. Jörg, T. Markus, B. Alice, "Non-Steroidal Anti-Inflammatory Drugs and Gabapentinoids for Chronic lumbar Pain: A Bayesian Network Meta-Analysis of Randomized Controlled Trials", *British Medical Bulletin*, vol. 01, pp. 85–95, 2021. https://doi.org/10.1093/bmb/ldab003

[9] H. Soltanali, M. Khojastehpour, J. T. Farinha, et al. "An Integrated Fuzzy Fault Tree Model with Bayesian Network-Based Maintenance Optimization of Complex Equipment in Automotive Manufacturing", *Energies*, vol. 22, pp. 1–21, 2021. https://doi.org/10.20944/preprints202108.0501.v1

[10] S. Zhao, Y. Shan, I. Gultepe, "Prediction of Visibility in the Arctic Based on Dynamic Bayesian Network Analysis", *Acta Oceanologica Sinica*, vol. 41, pp. 57–67, 2022. https://doi.org/10.1007/s13131-021-1826-z

[11] Zhou Lan, Yize Zhao, Jian Kang, Tianwei Yu, "Bayesian Network Feature Finder (BANFF): an R Package for Gene Network Feature Selection," *Bioinformatics*, vol. 32, no. 23, PP. 3685–3687, 2016. https://doi.org/10.1093/bioinformatics/btw522

[12] Z. M. Geng, Z. Q. Cai, Z. Zhang, et al. "Estimating Survival Benefit of Adjuvant Therapy Based on a Bayesian Network Prediction Model in Curatively Resected Advanced Gallbladder Adenocarcinoma", *World Journal of Gastroenterology*, vol. 25, pp. 5655–5666, 2019. https://doi.org/10.3748/wjg.v25.i37.5655
[13] C. Boya, D. Bansal, S. Kanakagiri, et al. "Efficacy and Safety of Opioid Analgesics for the Management of Chronic Low Back Pain: An Evidence from Bayesian Network Meta-Analysis", *Pain Physician*, vol. 24, pp. 73–82, 2021.
[14] P. Kumari, B. Bhadriraju, Q. Wang, et al. "A Modified Bayesian Network to Handle Cyclic Loops in Root Cause Diagnosis of Process Faults in the Chemical Process Industry", *Journal of Process Control*, vol. 110, pp. 84–98, 2022. https://doi.org/10.1016/j.jprocont.2021.12.011
[15] J. Xue, T. L. Yip, B. Wu, et al. "A Novel Fuzzy Bayesian Network-Based MADM Model for Offshore Wind Turbine Selection in Busy Waterways: An Application to a Case In China", *Renewable Energy*, vol. 172, pp. 897–917, 2021. https://doi.org/10.1016/j.renene.2021.03.084
[16] Z. Qi, K. Yue, L. Duan, et al. "Dynamic Embeddings for Efficient Parameter Learning of Bayesian Network with Multiple Latent Variables", *Information Sciences*, vol. 590, pp. 198–216, 2022. https://doi.org/10.1016/j.ins.2022.01.020
[17] S. Noriaki, T. Yoshinori, Y. Guangchuang, et al. "CBNplot: Bayesian Network Plots for Enrichment Analysis, *Bioinformatics*, vol. 10, pp. 2959–2960, 2022. https://doi.org/10.1093/bioinformatics/btac175
[18] W. Nguyen, P. O. Dusadeerungsikul, S. Y. Nof, "Plant Stress Propagation Detection and Monitoring with Disruption Propagation Network Modelling and Bayesian Network Inference." *International Journal of Production Research*, vol. 60, pp. 723–741, 2022. https://doi.org/10.1080/00207543.2021.2009139
[19] Z. Chen, Y. Lu, B. Cao, et al. "Driver Gene Detection Through Bayesian Network Integration of Mutation and Expression Profiles." *Bioinformatics*, vol. 10, pp. 2781–2790, 2022. https://doi.org/10.1093/bioinformatics/btac203
[20] T. Druet, M. Gautier, "A Hidden Markov Model to Estimate Homozygous-By-Descent Probabilities Associated with Nested Layers of Ancestors", *Theoretical Population Biology*, vol. 145, pp. 38–51, 2022. https://doi.org/10.1016/j.tpb.2022.03.001
[21] J. Li, X. Chen, Z. Li, "Spike Detection and Spike Sorting with a Hidden Markov Model Improves Offline Decoding of Motor Cortical Recordings", *Journal of Neural Engineering*, vol. 16, pp. 194–212, 2019. https://doi.org/10.1088/1741-2552/aaeaae
[22] L. Mao, W. Zhang, "Analysis of Entrepreneurship Education in Colleges and Based on Improved Decision Tree Algorithm and Fuzzy Mathematics", *Journal of Intelligent and Fuzzy Systems*, vol. 40, pp. 2095–2107, 2021. https://doi.org/10.3233/JIFS-189210
[23] J. R. Correa, M. Romero, "On the Asymptotic Behavior of the Expectation of the Maximum of I.i.d. Random Variables", *Operations Research Letters*, vol. 49, pp. 785–786, 2021. https://doi.org/10.1016/j.orl.2021.08.009
[24] A. Marshall, S. Lane, R. Ruiz, et al. "Online Discussion Blog Assignments: Design Implications for Teacher Education involving Mathematics", *Journal of Computers in Mathematics and Science Teaching*, vol. 38, pp. 321–335, 2019.
[25] M. V. Vien, J. Ai, C. K. Sung, "The Challenges of Implementing Information and Communications Technology (ICT) Based Online Learning in Chinese Independent High Schools (CIHS) in Malaysia", *Research in World Economy*, vol. 10, pp. 117–128. https://doi.org/10.5430/rwe.v10n2p117

7 Authors

Tan Feng obtained her MA in English Language and literature (2016) from SISU, Chongqing. Presently, she is working as a lecturer in the School of Humanities, Geely University of China, Jianyang. She has published several articles on CNKI and books as an associate editor and co-author. Her areas of interest include English teaching and English literature.

Laith Abualigah is an Associate Professor at Hourani Center for Applied Scientific Research, Al-Ahliyya Amman University, Jordan. He is also a distinguished researcher at the School of Computer Science, Universiti Sains Malaysia, Malaysia. He received his first degree from Al-Albayt University, Computer Information System, Jordan, in 2011. He earned a Master's degree from Al-Albayt University, Computer Science, Jordan, in 2014. He received a Ph.D. degree from the School of Computer Science in Universiti Sains Malaysia (USM), Malaysia, in 2018. According to the report published by Stanford University in 2020, Abualigah is one of the 2% influential scholars, which depicts the 100,000 top scientists in the world. Abualigah has published more than 250 journal papers and books, which collectively have been cited more than 9000 times (H-index = 45). His main research interests focus on Arithmetic Optimization Algorithm (AOA), Bio-inspired Computing, Nature-inspired Computing, Swarm Intelligence, Artificial Intelligence, Meta-heuristic Modeling, and Optimization Algorithms, Evolutionary Computations, Information Retrieval, Text clustering, Feature Selection, Combinatorial Problems, Optimization, Advanced Machine Learning, Big data, and Natural Language Processing. Abualigah currently serves as an associate editor of the Journal of Cluster Computing (Springer), the Journal of Soft Computing (Springer), and Journal of King Saud University – Computer and Information Sciences (Elsevier). E-mail: aligah.2020@gmail.com

Article submitted 2022-09-23. Resubmitted 2022-11-15. Final acceptance 2022-11-15. Final version published as submitted by the authors.