Exercise Reference Approach Based on Machine Learning based Linear regression model

[1]ANUSHA NEELAM, [2]M. ANJANEYULU

[1]PG Scholar, Dept. of MCA, Newton's Institute of Engineering, Macherla Guntur, (A.P)

[2]Assistant Professor, Dept. of CSE, Newton's Institute of Engineering, Macherla Guntur, (A.P)

Abstract: *In this study, we describe a machine-learning approach to the age-old problem of exercise suggestion. Based on the group to which a student belongs, this technique might suggest more relevant practise problems. To begin, we construct an accurate model of the students' understanding of each topic using linear regression and the EM method. Students are placed into one of three groups for each knowledge point based on their level of proficiency with that knowledge point and their overall level of competence. Determine the practise that may lead each kind of learner to a larger promotion for each knowledge point, based on their response history record. We utilise the k-nearest neighbour technique to categorise pupils, then offer activities that are tailored to each student's skill level and include the necessary knowledge points. Experiments have shown that this approach helps students progress more quickly while doing the same number of activities.*

Keywords: *Exercise recommendation, Machine learning, Linear regression, EM algorithm, K-means, K-NN.*

I. INTRODUCTION

The allocated pupils will make more progress in the same amount of time if they are given tasks that are more suited to their abilities. From among the numerous exercises that include the same knowledge points, the current exercises recommendation algorithm is unable to choose the activities that are appropriate for the intended pupils. This research provides a problem suggestion approach based on machine learning to address the drawbacks of current workout recommendation systems. The following are the paper's contributions:

(1) We classify students into groups according to their mastery of each topic and then suggest activities tailored to those groups using k-means and k-NN algorithms.

(2) Compared to students who do the same number of tasks as we suggest, our students show greater gains in information retention and application. We addressed

ISSN: 2366-1313

the issue of the current exercise recommendation technique not being able to ensure that the suggested exercises would result in a better development in the student's abilities than alternative activities.

II. RELATED WORKS

Weak knowledge point exercise suggestion techniques, exercise recommendation methods based on goal difficulty, and exercise recommendation methods designed to boost student performance on a specific test are the two most common kinds of exercise recommendation methods used today.

Students may be guided towards strengthening their weaker areas of knowledge with the use of an exercise suggestion approach that takes goal knowledge points into account. A suggestion approach was presented by Y. Huo et al. [1] to aid students in identifying knowledge gaps and receiving recommendations for corresponding practise problems. L. Fang developed a strategy for recommending tasks to pupils [2] that would help them practise English grammar with a certain structure in mind.

Yan Y suggested a strategy for recommending coding activities in C to pupils. In order to help students, strengthen their knowledge, this technique may identify the specific modules in which they are deficient [3].

The recommended workouts for pupils may be tailored to their own skill levels using the target difficulty exercise suggestion approach. Medium-difficulty tasks were provided by Changning J et al. [4] and might be recommended for target students.

Students may benefit from the issue suggestion approach, which was developed with the goal of raising their test scores on a certain exam. Iwata T suggested a strategy for recommending exercises that would improve students' performance on a targeted English test using a logistic regression model [5]. A technique for recommending practise drills in an individualised learning environment was suggested by Pavel Michlik.

The purpose of this strategy is to improve kids' test scores [6].

III. METHOD DESIGN

This strategy begins by classifying learners into subgroups based on their current level of understanding of each topic, then determines which exercises are most suited to learners at each of those levels, and lastly determines which subgroup the intended audience of learners falls into and provides recommendations accordingly. Here are the detailed procedures involved in this approach:

ISSN: 2366-1313

(1) We begin by using the procedure outlined in 3-A to determine the extent to which each student has mastered each topic.

(2) We classify each student into one of five categories, representing the top 20%, the next 40%, the next 60%, the next 80%, and the final 100%, based on how well they have mastered the related body of information. The students were then categorised into groups based on how they performed on every test that included a knowledge point (for details, see Section 3-B).

(3) We utilise the procedure outlined in 3-C to determine which exercises are most appropriate for each group of pupils and then propose just those.

(4) Based on the topics that students have indicated they want to learn more about, we will suggest tasks that include those topics. First, we determine which category the target students fall into based on the knowledge points in question and the extent to which the target students already know those points. The k-NN technique (detailed in Section 3-D) is then used to determine the subset of these students who fall into the target category. Finally, we advise pupils to do the related activities for this group.

A. The Method for Determining How Much Students Have Learned

First, we collect all of the students' past response data, which includes things like their identifiers, the exercises they've completed, the amount of knowledge points each activity tested, their average score on those exercises, and their actual performance on the exercises.

The term "knowledge points mastery index" is introduced.

The idea of knowledge point mastery index is a way to quantify how much lower or higher than average a person's mastery of a knowledge point is relative to the norm.

When the value of a knowledge point is unknown or known, we provided the formulae for determining a mastery index based on the number of knowledge points mastered.

Without specifying a weight for a knowledge point, the following formula may be used to determine the knowledge point mastery index:

$$G_{SK} = \frac{\sum_{i=1}^{n}(s_i - m_i)}{\sum_{i=1}^{n} t_i} \qquad (1)$$

Among them are the mean score of exercise, the total score of exercise, and the score that student S got on an exercise involving knowledge point K that the student completed.

Under the assumption that a weight has been given to each knowledge point, the following equations may be used to

ISSN: 2366-1313

determine the knowledge point mastery index:

$$h_i = \sum_{j=1}^{n} k_j x_j + b \qquad (3)$$

Among them are the mean score of exercise i, the weight of knowledge point K in exercise i, the overall score of exercise i, and student S's score on an exercise using knowledge point K that S has completed.

The following procedures are used to determine a student's level of understanding based on their answers in the past;

We start by selecting questions with a single knowledge point, then we utilise formula (1) to approximately compute and record each student's knowledge point mastery Index.

To determine how much emphasis should be placed on each piece of information in the questions, we use the linear regression model. Using Exercise E and Student S as examples, we demonstrate how to use the linear regression model. The following is a formula for predicting student s's score on Exercise e, assuming there are no knowledge points.

$$h_i = \sum_{j=1}^{n} k_j x_j + b \qquad (3)$$

is the weight of, is the expected score on Exercise E for Student S, and is the student's Index of Mastery of Exercise E's jth Knowledge Point. After training the model with all of the students' past response data, we can determine how much importance each knowledge point in each exercise really has.

When the relative importance of each knowledge point is understood, we may utilise formula(2) to more accurately determine which students have mastered which information points.

The EM method is used to determine the present value of the students' knowledge point mastery Index, which is then used as the basis for future assessments. The relative importance of each knowledge point in subsequent exercises is determined using linear regression and the most recent value of the Knowledge Point Mastery Index. Then, the new group of students' knowledge point mastery Index is determined using formula (2) and the weight of knowledge point in the new set of tasks.

The ultimate result of the Knowledge Point Mastery Index was acquired after 100 rounds of iteration.

Grouping the Students, Part B

As an example, we'll use knowledge point K to show how to cluster a class. The three possible states of a response to a question are "correct," "incorrect," and "not applicable." We use students' responses to exercises testing their understanding of knowledge point K as the cluster features.

ISSN: 2366-1313

To achieve this grouping of pupils, we use the k-means method.

Choice of Suggested Workouts, Section C

To determine how well an exercise is at helping students improve their mastery of a particular knowledge point, we first define the concept of knowledge point mastery gap, which is used to express the change in students' mastery of the specified knowledge point between the beginning and end of an exercise. Student S's pre- and post-exercise K knowledge point mastery gap may be calculated using the following formula.

$$D = Gsk2 - Gsk1 \qquad (4)$$

before doing Exercise E, make a dent in Knowledge Point S.

Using the procedure outlined in 3-A, we can determine Student S's Knowledge Points Mastery Index based on the five Knowledge Point K tasks she completed prior to Exercise E. Student S's mastery index in knowledge point K, after completing exercise E, is denoted by Gsk2.

Using the procedure outlined in 3-A, we can determine Student S's Knowledge Points Mastery Index based on the five Knowledge Point K activities she completed after Exercise E.

We determine each student's knowledge point mastery gap for all exercises they've completed for their knowledge point and their student category. When calculating the average knowledge point mastery gap for an activity, we first identify all students in the allocated group who have completed the exercise in question. Finally, for each knowledge point and student group, we identify the 10 activities with the largest average mastery gap as the best options for that group.

D. Students' Evaluations of Different Groups

As an example, we use knowledge point K to classify students into one of many groups. The k-means method was used to classify the pool of target students into 8 groups before the categories were judged. To determine which group each student belongs to, we use the scikit-learn function corresponding to the k-NN method with the default values for the algorithm's parameters and the students' responses to all exercises including knowledge point K as the characteristic.

IV. EXPERIMENT

A. Experimental Setting

The goal of this study is to determine whether or not the recommended exercises actually help students' abilities more than the regular exercises, and to draw conclusions about which types of exercises are most effective at bringing about these improvements.

ISSN: 2366-1313

An online education firm generously supplied the data we utilised in our investigation. The information covers the whole city's first-year junior high school student body and their online responses to a specific mathematics chapter between May 1, 2020, and May 15, 2020. Each of the 130k data points comprises the student's id, the exercise id, the question's score, the student's actual score, and the time at which they began responding questions. 1,682 student answers to 421 different activities were utilised to generate the data for this study.

Experimentation Methodology

The other half of the class was used as a test group. We utilise the approach described in this work to categorise the students in the training set and identify the appropriate activities for each learner type for each knowledge point. To conduct the experiment, we chose five areas of knowledge and randomly picked one hundred students from the remaining group to serve as the test set. We utilise the k-NN algorithm to categorise each test subject based on their knowledge of these five topics. For each knowledge point and test subject, we chose at random two exercises that the subject had completed that included the knowledge point: a "recommended" activity and a "general" exercise.

We have computed the average knowledge point mastery gap of all suggested tasks completed by all students, and this result is recorded as 1 for each of the five knowledge points we chose. We have estimated the average knowledge point mastery gap of all the general exercises that all students have done (indicated as the average value 2) for each of the five knowledge points we picked. The number of students whose suggested activities had a knowledge point mastery gap larger than the value of the general exercises (marked as 1) was tallied for each of the five knowledge points we chose. The number of students whose suggested exercise knowledge point mastery gap was less than or equal to the values of the general exercises, which were indicated as number 2, was tallied for each of the five knowledge points we chose.

C. Methods of Experimentation

The average value of 1 for each knowledge point in this experiment is more than the average value of 2 for each knowledge point, as shown in table 1, and the number 1 for each knowledge point is bigger than the number 2. What this indicates is that compared to students who did not undertake the activities indicated by this

ISSN: 2366-1313

approach, those who did so had a larger improvement in their information retention after completing the prescribed exercises. Inferring from this, pupils may make more progress in the same amount of practise time with this approach.

Table 1 Student statistics table

Knowledge point	Average value 1	Average value	Number 1	Number 2
Knowledge point 1	0.13	0.15	61	39
Knowledge point 2	0.19	0.21	64	36
Knowledge point 3	0.09	0.10	53	47
Knowledge point 4	0.22	0.24	54	46
Knowledge point 5	0.15	0.18	62	38
Total	0.16	0.18	294	206

VI Conclusion

This research puts out an AI-based strategy for advising on physical activity. Based on the student group, this approach may suggest activities that will have the greatest impact on skill development. Experiments show that using this strategy leads to larger gains for pupils in the same amount of practise.

REFERENCES

[1] J. Xia, G. Li and Z. Cao, "Personalized exercise recommendation algorithm combining learning objective and assignment feedback," Journal of Intelligent & Fuzzy Systems, vol. 35, (3), pp. 2965- 2973, 2018.

[2] Y. Huo, J. Xiao and L. M. Ni, "Towards personalized learning through class contextual factors-based exercise recommendation," in 2018, . DOI: 10.1109/PADSW.2018.8644555.

[3] Y. Yan et al, "A method for personalized C programming learning contents recommendation to enhance traditional instruction," in 2017, . DOI: 10.1109/AINA.2017.13.

[4] Changmeng J , Jun F , Xia S , et al. Personalized exercises Recommendation algorithm based on Knowledge Hierarchical Graph, ReKHG[J]. Computer Engineering and Applications, 2018.

[5] Iwata T , Kojiri T , Yamada T , et al. Recommendation Algorithm for Learning Materials That Maximizes Expected Test Scores[C]// Pacific Rim International Conference on Artificial Intelligence. Springer, Berlin, Heidelberg, 2008.

[6] Naga Lakshmi Somu, Prasadu Peddi (2021), An Analysis Of Edge-Cloud Computing Networks For Computation Offloading, Webology (ISSN: 1735-188X), Volume 18, Number 6, pp 7983-7994.

[7] Pavel Michlík, Mária Bieliková. Exercises recommending for limited time learning[J]. Procedia Computer Science, 2010, 1(2):2821-2828