Prediction Method of Online Teaching Effect of Career Guidance Course Based on Multi Task Learning

Wei Han(✉) and Mingsheng Li

Baotou Teachers' College, Baotou 014030, China
coolhanwei@163.com

Abstract. There are issues with intermittent and lagging schedules in the teaching of employment guidance courses, as well as a lack of professional teachers. Online teaching can solve these problems. Therefore, in order to better analyze the learning effectiveness of online teaching in employment guidance courses, a multi task learning based prediction method for the learning effectiveness of online teaching in employment guidance courses is designed to determine whether learners have mastered the knowledge content of the employment guidance courses they are learning. Design a Chinese word segmentation sequence annotation method based on multitasking learning, implement preprocessing of online teaching data for employment guidance courses, and improve the generalization ability and effectiveness of the model. For the data after Chinese word segmentation, an improved K-means algorithm is used to mine the learning behavior information of students, thereby discovering potential patterns and patterns hidden in the data. Propose the EduHawkes model, which adds various types of feature information to predict learning outcomes based on analyzing student behavior patterns, achieving online teaching of employment guidance courses and improving the accuracy and interpretability of learning outcomes prediction. The test method shows that the accuracy of the method and the Matthews correlation coefficient are both high.

Keywords: Multi Task Learning · Online Teaching of Career Guidance Courses · K-Means Algorithm · Chinese Word Segmentation · Learning Effect Prediction

1 Introduction

If colleges and universities want to ensure their overall sustainable development, they must solve the employment problem of graduates. Only when the school successfully sends millions of graduates to jobs can it build the cornerstone of a harmonious and stable society [1]. Therefore, all colleges and universities should regard employment guidance as a major event. In terms of the current situation of employment guidance courses in China, there are still the following problems:

There are discontinuities and delays in the time arrangement of career guidance courses. Employment guidance courses in colleges and universities are generally offered

G. Gui et al. (Eds.): eLEOT 2023, LNICST 543, pp. 17–32, 2024.
https://doi.org/10.1007/978-3-031-51465-4_2

in the form of public elective courses. Although the courses involve the teaching arrangement of career planning, employment skills, employment policies and other contents, due to the lack of attention to employment guidance courses in colleges and universities, students do not have a deep understanding of the course. In addition, students are under great pressure to learn professional courses, and often only choose employment skills employment policy and other courses, while career planning and other knowledge are often ignored. The teaching effect of employment guidance can not be shown in students' career development, and guidance only stays on the surface. Although the Ministry of Education of the People's Republic of China requires that career guidance and career development courses be offered in the form of compulsory courses, until now, most colleges and universities have changed their career guidance courses to compulsory courses, but they are still optional. In addition, the career guidance courses of many colleges and universities only carry out several career lectures in the graduating students, without systematic and complete career guidance teaching training.

The teaching content of employment guidance lacks professional pertinence and the teaching form is single.

On the one hand, the content of textbooks is universal; On the other hand, the content of employment is mostly theoretical and has little practical significance, such as employment information, employment skills, employment system, etc. At present, the career guidance courses in many colleges and universities are still dominated by classroom teaching, lacking diversified teaching methods, which tend to ignore the different needs of students for the courses, and it is difficult to mobilize learning interests, which is not practical.

Employment guidance teachers lack professional teachers.

Teachers of employment guidance and career planning should not only have basic knowledge of employment, occupation, psychology, law, etc., but also have professional knowledge, be familiar with the characteristics of different majors, and master various policies and regulations in related fields. At present, the main teachers of employment guidance courses in most colleges and universities are counselors or staff of employment guidance centers. The lack of career guidance instructors and professional field instructors can not accurately grasp the employment situation of various industries, and the guidance ability is limited, so the employment guidance effect of colleges and universities is not significant [2].

In order to solve the above problems, online teaching methods of career guidance courses have emerged. In online teaching of career guidance courses, we need to pay attention to an important issue, that is, how to judge whether a learner has mastered the knowledge he is learning. Based on this background, this paper studies the prediction of online teaching learning effect of career guidance courses.

For the research on the prediction of learning effect, the current research results have been relatively rich. Some scholars have proposed a prediction method of students' online learning effect based on rank correlation analysis: for the problem that students have many online learning behavior characteristics and the accuracy of learning effect prediction is not high. A feature selection method based on rank correlation is proposed. By calculating the rank between each behavior feature and learning effect, the correlation degree of each feature and learning effect is obtained. On this basis, multiple features with the highest correlation degree are selected to form a new data set, and then normalization and one pot coding are completed, and machine learning methods and neural network models are used for training respectively. Finally, the online learning effect prediction of students is realized. On the publicly published learning data set, the rank correlation is used for feature selection combined with the multi-layer perceptron model, and the learning effect of students can be correctly classified, and the F value is higher than the traditional linear correlation selection method. Other scholars have proposed a learning process analysis and learning effect prediction method based on behavior sequence: learning process data reflects the state of learners in the learning process. Currently, many researches have been conducted on data mining and analysis of learners' learning process. Most of these coarse-grained data are based on the energy and time invested by learners in a certain learning behavior, which can not reflect the level of learners' cognitive investment in detail, and some learning behavior data have low accuracy in predicting learning effects. Compared with learner participation, the learning behavior sequence in the learning process can better reflect the learning behavior trajectory, willingness and cognitive process of learners. The lag sequence analysis method is used to analyze the learning process data on the DEEDS platform. It is found that the lag sequence analysis method can clearly reveal the important learning behavior sequence; Compared with support vector machine, logistic regression, decision tree and other data mining methods, naive Bayesian method has good prediction performance, with an average accuracy of more than 70%. The research results prove that learners' learning behavior sequence can provide teachers with a more comprehensive online learning picture to help teachers discover learners' learning habits, preferences and cognitive processes, assist teachers to reflect on the teaching process, and accurately predict learners' learning achievements through behavior sequence data, and then analyze the key attributes in the prediction model. To provide suggestions for teachers to take targeted intervention measures in the follow-up teaching process, so as to improve the educational and teaching performance. The above methods exist.

This paper designs a method to predict the learning effect of online teaching of career guidance courses based on multi task learning.

2 Design of Online Teaching Learning Effect Prediction Method for Career Guidance Course

2.1 Chinese Word Segmentation Sequence Labeling

Based on multi task learning, a Chinese word segmentation sequence labeling method is designed to implement online teaching data preprocessing of career guidance courses. This method uses tag consistency mechanism based on multi task learning to conduct in-depth training and learning, which is divided into three modules, namely, pre tagging module, multi task shared learning module, and joint training module [3].

The way of multi task learning can avoid the further spread of errors, and combine multiple tasks to learn together, making the information exchange between multiple tasks more convenient, and improving the accuracy and learning efficiency of each sub task. However, in the previous study, we mentioned that most of the existing Chinese word segmentation methods based on multi task learning are based on different data sets of different tasks, which can only ensure that multiple tasks in form learn together. In essence, each task is still relatively independent, which will inevitably lead to information conflicts, excessive noise Information sharing is limited, and very important boundary information cannot be learned, which affects the final learning effect. In order to solve the above problems and make it possible to learn tag consistency information, this study adds a pre labeling module to enhance the learning of available labeled resources and alleviate the pressure of data shortage.

First, the original named entity dataset, word segmentation dataset, and part of speech tagging dataset are recorded as:

$$M_{ner} = \left\{ m_j^{ner} \right\} \tag{1}$$

$$M_{cws} = \left\{ m_j^{cws} \right\} \tag{2}$$

$$M_{pos} = \left\{ m_j^{pos} \right\} \tag{3}$$

among m_j^{ner} Represents the number of j Samples; m_j^{cws} Represents the No j Samples; m_j^{pos} Represents the number of POS data sets j Samples.

Then, we use the M_{cws}, M_{pos} Train the corresponding word segmentation device and part of speech tagging device. Next, the benchmark data set M_{ner} The original samples in are de labeled and processed into plain text format. The trained word segmentation device and part of speech tagging device are used to perform word segmentation and part of speech tagging operations based on word granularity. Since the four word tagging method is used in the experiment, the four word segmentation tagging based on word granularity is required for the results obtained. Since the named entity data set uses a three word bit {B, I, O} tagging system, in order to achieve tag alignment, the four word bit word segmentation is labeled {B, M, E, S}, which is also converted into {B, I, E, S}, and the part of speech tag is also processed accordingly. For example, if the tag representing the part of speech of a noun is n, then the tag of a two word noun is {B-n I-n}, which is split and mapped to each word [4].

Finally, extract the final obtained word segmentation and part of speech tags, integrate the tags and align them to the benchmark NER dataset, and then obtain a special multi task learning dataset that neatly corresponds to the three task tags at the same time. As shown in the example in Table 8, in case of ambiguity in the pre labeling process, the label of the datum named entity identification data set shall prevail for alignment. For example, the person name entity <Wang Xiaoming> may get the segmentation of <Wang|Xiao|Ming> and the word segmentation label of <B B B> after word segmentation, which does not correspond to the original NER label <B-PER I-PER I-PER> on the border. In this case, we modify the word segmentation label to <BIE> , that is, the entity boundary shall prevail.

After the above processing, we get that the pre labeled dataset is recorded as:

$$S = \{s_i\} \tag{4}$$

One sample s It is composed of character set, character named entity recognition label, word segmentation label and part of speech label.

On the basis of the pre labeling module, a multi task shared learning module is constructed, that is, joint learning for multiple tasks. The multi task shared learning module includes two parts: the shared presentation layer and the sequence decoding layer. In the following description, the input sequence is recorded as:

$$B = \{b_1, b_2, \ldots, b_n\} \tag{5}$$

among b_n Representation sequence B Of n Words. B Represents the input of a special data set D that is pre labeled.

(1) Shared presentation layer
1) BERT layer

Traditional static word vectors are usually unable to represent polysemy of a word. In this stage, a pre training language model based on BERT is selected for feature learning to represent context sensitive semantics. For the input sequence, the sequence is uniformly processed into the form that the start token is the special category embed character "[CLS]" and the end is the special separator character "[SEP]" for encoding. BERT uses the self attention mechanism to learn the dependency relationship and context semantics of each word with other words, and then uses feedforward neural networks to perform projection transformation of multiple different linear transformation pairs on the input after the Attention calculation, and finally obtains the overall information of the sequence. After the BERT layer coding calculation, the input sequence will be B The conversion is recorded as:

$$F = \{f_1, f_2, \ldots, f_n\} \tag{6}$$

2) BiGRU layer

On the basis of the BERT layer, we add a two-way gating cycle unit layer (BiGRU layer) to further learn about information sharing between multiple tasks. Output sequence of BERT layer $F = \{f_1, f_2, \ldots, f_n\}$ As the input of the BiGRU layer, for the input of the time r process, the update gate zrand reset gate rrjointly determine the output of the hidden state. Record the output of two-way GRU forward and backward hidden layers as J_{front}, J_{back}, remember that the final output vector is J.

(2) Sequence decoding layer

On top of the shared presentation layer, we will build a sequence decoding layer according to the characteristics of the task, and decode for different tasks. In order to model the adjacency between tags, conditional random field is selected as the decoding model.

This paper proposes a tag consistency mechanism throughout the whole training process based on multi task learning Chinese word segmentation sequence tagging method. In the joint training module, the specific application of tag consistency mechanism is reflected in the loss calculation. For the problem of whether the decoding tags of different tasks are consistent, a more universal method [5] is defined.

Specifically, we first defined the loss function as shown in Formula (7):

$$loss = \omega_1 * l_{cws} + \omega_2 * l_{pos} + \omega_3 * \mathrm{l}_{ner} + \omega_4 * \mathrm{l}_{con} \tag{7}$$

In Formula (7) l_{cws}, l_{pos}, l_{ner} The losses of word segmentation task, named entity recognition task and part of speech tagging task are calculated respectively based on the standard (ground truth) label of their respective tasks;l_{con} It is the supplementary loss defined according to the tag consistency mechanism, representing the loss of cross task tag consistency; $\omega_1, \omega_2, \omega_3, \omega_4$ representative l_{cws}, l_{pos}, l_{ner}, l_{con} Weight of.

By optimizing this loss, the model can grasp more boundary information and learn the consistent content of multi task decoding tags, such as the alignment between the boundary of word segmentation tags and the boundary of entity tags and part of speech tags, which to some extent alleviates the problem of non-standard definition of word segmentation standards. On the basis of special data sets. The network is adjusted and updated through the calculation of loss, which further emphasizes the consistency and reduces the error caused by the noise of multi task learning, so as to realize the effective sharing and fusion of multiple information under the multi task learning mode.

2.2 Information Mining of Students' Learning Behavior

The improved K-means algorithm is used to mine the information of students' learning behavior from the online teaching data of career guidance courses labeled with Chinese word segmentation sequence.

As a partition based clustering algorithm, K-means algorithm is widely used for its advantages of easy understanding, easy implementation, and strong interpretability. However, K-means algorithm also has some shortcomings, such as the number of clusters K value needs to be manually set, too large or too small value of K will affect the clustering effect, the initial cluster center is randomly determined, more sensitive to isolated points, etc. These shortcomings will affect the final effect of clustering, resulting in poor stability and quality of clustering results. The K-means clustering algorithm is optimized and improved from multiple factors. The main idea of the heuristic method is to solve how to deal with the next step according to the empirical rules. In the iterative calculation process of clustering, this study uses the heuristic method to reduce the calculation time required for sample point allocation and improve the operation efficiency. In each iteration process, a simple data structure is used to store the distance between each point and its nearest cluster. In the next iteration process, first calculate the distance between each point and its previous nearest cluster. If the new distance obtained by calculation is less than or equal to the previously saved distance, then the point will remain in the original cluster, and there is no need to calculate the distance between the point and other clusters. This saves the calculation time [6] of the distance between this point and other clusters.

Aiming at the problem that K-means algorithm is sensitive to outliers, this research first optimizes the iterative calculation process of K-means++ algorithm based on heuristic method, which improves the calculation speed of K-means++ algorithm, and then combines the optimized K-means++ algorithm with LOF method to improve the operation efficiency of outlier detection algorithm.

The main idea of the maximum minimum distance algorithm is to set the distance threshold, and then determine the initial cluster center. The maximum and minimum distance algorithm is used to determine the initial cluster center, as shown in Formula (8).

$$\begin{cases} h = \max\{w(i,j)\} \\ i = 1, 2, \ldots, p \\ j = 1, 2, \ldots, q \end{cases} \tag{8}$$

among $w(i,j)$ Represents a data object x_i The minimum distance from the existing initial center point; p Remove the number of samples selected as the initial cluster center from the dataset; q The number of initial cluster centers selected for yourself.

In K-means algorithm, if a very isolated sample point is determined as the initial cluster center, poor clustering results will be obtained.

In this study, the idea of maximum and minimum distance is combined with the improved outlier detection algorithm to select the initial cluster center. First, the optimized outlier detection algorithm is used to detect outliers, and at the same time, the candidate set of outliers and the set of outliers are obtained. The set of outliers is deleted from the original dataset and stored temporarily to avoid the impact of outliers on the final clustering results. Then, based on the candidate set of outliers, select the first two initial cluster centers, calculate the distance between all the points in the candidate set of outliers, and find the two points that are farthest away as the initial cluster centers.

Based on the above improvement ideas, this paper first determines the number of clusters by combining the elbow rule and the contour coefficient method to avoid the impact of artificially specified cluster numbers on the clustering results, and then conducts outlier detection based on the improved outlier detection algorithm to avoid the adverse impact of outliers on the final clustering results. At the same time, the initial cluster center is selected step by step based on the candidate set of outliers and the idea of maximum and minimum distance, avoiding the impact of random selection of initial cluster center on the clustering results, and ensuring the stability of clustering results and the quality of clustering results. The improved clustering algorithm is as follows:

Input: set E of n data points.

Output: clustering result cluster.

1. Elbow rule and contour coefficient method determine the clustering number U of dataset E1;
2. Use the improved outlier detection algorithm to detect outliers, and obtain "outlier candidate set" E2, "outlier set" E3;
3. Obtain U initial cluster centers Centroid in E2, E3, E1 according to the improved method;
4. For each point ejin E, find the centroid cj ∈ Centroid nearest to it, and assign ejto cluster j;
5. Order:

$$\begin{cases} ClusterId[i] = j \\ NearestDist[i] = w\left(e_j, c_j\right) \end{cases} \tag{9}$$

6. While does not meet the convergence condition
7. For each cluster ClusterDict, calculate the new centroid CentroidDict;
8. Reassign each sample point to the nearest cluster;
9. End

Although the improved clustering algorithm spends extra time in selecting the initial cluster center, the number of iterations is reduced, and the calculation time of the iterative process is also reduced. In terms of running time, for small data sets, the improved algorithm is basically consistent with the original K-means algorithm, or reduced; For large data sets, it is slightly higher than the original K-means algorithm, but the optimized clustering algorithm not only has a very stable clustering result each time, but also has a better clustering effect [7].

2.3 Design of Learning Effect Prediction Model

The current research mostly uses exam scores as the standard to evaluate students' learning effect, but the impact of exam scores on learning effect in online learning is limited. Students obtain knowledge by watching videos on online learning platforms, resulting in a large number of learning behaviors being ignored by existing research methods. At the same time, the current research lacks the description of other characteristics, which limits the effectiveness of the prediction process. In view of these two problems, EduHawkes model is proposed. Based on the analysis of student behavior patterns, various types of characteristic information are added to predict the learning effect, so as to realize the prediction of online teaching learning effect of career guidance courses.

Firstly, the learning behavior sequence is input into the new Hawkes process to obtain the learning behavior representation vector. Secondly, the student and curriculum feature vectors are used as input, and the fusion feature vector is obtained through the counter position multiplication between the feature vectors. Finally, the fusion feature vector and the learning behavior representation vector are input into the classification model for classification after the counter position addition operation.

EduHawkes model is shown in Fig. 1. The model consists of three parts: representational learning behavior sequence, feature fusion and classification prediction [8].

At present, the commonly used sequence representation methods are based on the most basic deep learning model, which has poor interpretability and will lose the information implied in the sequence. The new Hawkes process makes up for these two shortcomings of the current representation methods. It not only inherits the high interpretability of the Hawkes process, which can represent and model online learning modes, but also has the ability of deep learning and autonomous learning implicit features, bridging the gap between sequence understanding and prediction in traditional methods.

The learning behavior sequence is constructed based on the mining data, and then the unlabeled training set (that is, all the behavior events on the sequence) is trained through the new Hawkes process. Finally, the multi-dimensional learning behavior representation vector is obtained by inputting the learning behavior sequence and outputting it.

In reality, it often happens that different students make similar learning behaviors, but their learning effects are quite different. The cognitive state of students in learning

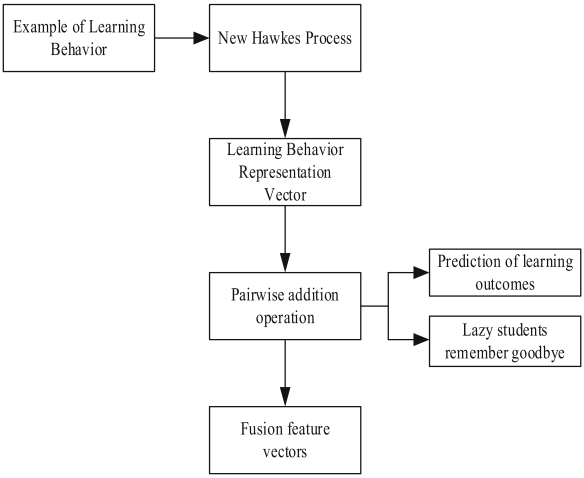

Fig. 1. EduHawkes Model

is restricted by many factors. Therefore, when the model is used to predict the learning effect, if it is only based on the data of students' learning behavior, the interpretability of the model will be poor. To solve this problem, this paper extracts student and curriculum features from the data set as new inputs to the model, so that the model can incorporate new features to obtain more ideal prediction results [9].

In the process of online learning, the interaction between students and online tutoring platform is the basic way of learning and the main form of student characteristics. According to the type of interaction, it can be divided into three categories: assessment interaction, learning content interaction, and interpersonal interaction. At the same time, attention is also one of the cognitive factors that affect learning. A large number of literatures have discussed the impact of the Internet environment on the attention mechanism, pointing out that attention has become a scarce resource in the Internet era full of information and information, which is easy to disperse and dissipate, thus causing cognitive problems. The impact of the Internet environment on learning attention investment has gradually attracted attention in the education field. The attention situation of online learning restricts the online learning effect of students. Therefore, when selecting students' characteristics through data sets, this paper takes into account the characteristics related to students' attention while selecting interaction type characteristics. The classification description of students' main characteristics is shown in Table 1:

Although online tutoring platform has massive teaching resources, it also brings a series of problems, such as uneven course quality, which brings some problems to students when choosing courses and affects their learning efficiency. Therefore, when selecting course features through data sets, this paper selects the features most relevant to course quality, which are mainly divided into three categories: course attributes, course popularity, and course assessment. The main features of the course are shown in Table 2:

Table 1. Classification of main characteristics of students

S/N	Characteristic type	Characteristic description
1	Appraisal interaction	Number of exams
		Number of classroom quizzes
		Number of job submissions
2	Learning content interaction	Times of watching online video
3	Interpersonal interaction	Number of theme posts published
		Number of posts viewed
		Number of comment posts
4	Attention related	Preferred course type
		Students' knowledge background

Table 2. Classification of main characteristics of courses

S/N	Characteristic type	Characteristic description
1	Course Properties	Course category
2	Popularity of courses	Number of liked courses
		Number of course comments
		Cumulative number of viewers of the course
3	Course assessment	Completion rate of course examination
		Submission rate of course assignments

After selecting specific student and course features, the features need to be extracted from the corresponding data set. Unique heat coding is a common method of feature extraction. Calculate the total number of features N, and each feature can be represented by a vector consisting of N-1 zeros and a single 1. However, there are two problems in doing so: 1. If there are many features, the transformed vector will have a huge dimension and be too sparse. The performance of deep learning for sparse input is not good. 2. Since maps are independent of each other, they cannot represent the relationship between different features. Considering these two problems, the method of feature embedding is used to express each feature in vector form with less dimensions, and to express the relationship between features, so as to better represent the feature space. Feature fusion is composed of two parts, namely, student feature vector and course feature vector. The student characteristic matrix is as follows:

$$T = \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ t_{m1} & t_{m2} & \cdots & t_{mn} \end{bmatrix} \tag{10}$$

In Eq. (10) m Indicates the number of students, n Represents the dimension of student vector. Take out each row in the matrix, and the eigenvector of each student can be expressed as:

$$t = [t_{m1}, t_{m2}, \ldots, t_{mn}] \tag{11}$$

Similarly, the curriculum characteristic matrix can be expressed as:

$$D = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ d_{m1} & d_{m2} & \cdots & d_{mn} \end{bmatrix} \tag{12}$$

The eigenvector of each course can be expressed as:

$$d=[d_{m1},\ d_{m2},\ \ldots,\ d_{mn}] \tag{13}$$

Finally, according to the current behavior sequence, the corresponding student feature vector and course feature vector are taken out and multiplied to get the fusion feature vector, as shown below:

$$u_s = t \oplus d \tag{14}$$

As mentioned above, it is believed that the prediction results of students' learning effects will be affected by learning behavior patterns, student characteristics and curriculum characteristics. Based on this idea, after obtaining the representation vector of the learning behavior sequence using the new Hawkes process, select the fusion feature vector obtained with the above, and perform the addition operation to form a new vector as the input of the classification model. The feature operation process is shown in Fig. 2:

Before the learning effect prediction results are output through the classification model, the full connection layer needs to be used for processing. The full connection layer is a linear combination of data without considering the nonlinear nature of the activation function, which can also be called "linear layer". The formula is as follows:

$$Y = EX + g \tag{15}$$

In Eq. (15) Y Refers to the function dependent variable; X Refers to the function independent variable;E Means a linear function; g It refers to the offset vector with dimension n. This paper sets the dimension of n as 2 or 4 dimensions [10] for different classification tasks.

The output of the linear layer is sent into the softmax function for probabilistic classification, and the output of multiple neurons is mapped into the interval of (0,1), which is analyzed in the form of probability, so as to realize the classification task.

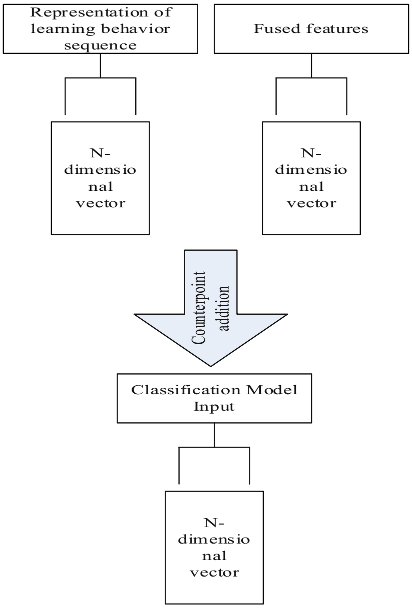

Fig. 2. Feature operation process

3 Example Application

3.1 Experimental Process

This paper tests the performance of the online teaching and learning effect prediction method of the career guidance course based on multi task learning.

Data preprocessing refers to a series of data processing before the experiment. At present, most of the data obtained through the teaching platform are incomplete, and there are inconsistent problems in the format, so it is impossible to directly input the model for experiments. There are many ways to preprocess: data cleaning, data reduction, data transformation, etc. The application of these processing technologies has greatly improved the quality of data and the accuracy of experimental results.

Data preprocessing has a great impact on the results of learning effect prediction, because the original dataset contains a lot of data unrelated to video viewing behavior, including the case of characters, spaces, special letters, etc. In order to reduce the impact of data on the experimental results and make the data more concise, the rules shown in Table 3 are used to process the data.

The data set is processed according to the above rules. Each behavioral data after processing is composed of the elements shown in Table 4, which are: user ID, video ID, behavioral type, timestamp, video start time and video end time.

First, the Chinese word segmentation sequence tagging method based on multi task learning is used to implement the preprocessing of Chinese word segmentation of experimental data. Then the improved K-means algorithm is used to mine the learning behavior information of students. Finally, through the designed EduHawkes model, on the basis

Table 3. Data Preprocessing Rules

S/N	Data preprocessing rules	Data preprocessing method
1	Unrelated data processing	Eliminate irrelevant behavior data
2	Special letter processing	Delete special letters contained in the dataset
3	Data segmentation	Split the log data

Table 4. Data Composition

S/N	label	explain
1	v_end_t	Video end time
2	v_start_t	Video Start Time
3	t_s	time stamp
4	stu_beh_t	Video viewing behavior type
5	video_id	Video ID
6	user_id	User ID

of the analysis of student behavior patterns, various types of feature information are added to predict the learning effect, so as to realize the prediction of the online teaching learning effect of career guidance courses.

The design method evaluates two key applications in the field of online learning:

(1) Evaluation of learning effect
(2) Slack student identification

Evaluate the performance of the model in learning effect evaluation task and slack student identification task through the following indicators:

(1) Accuracy
(2) Matthews correlation coefficient

3.2 Comparison Method

In the test, the online learning effect prediction method based on rank correlation analysis and the learning process analysis and learning effect prediction method based on behavior sequence are used as comparison methods, which are represented by method 1 and method 2 respectively.

3.3 Test Results

3.3.1 Accuracy Test Results

The design method and the accuracy test results of method 1 and method 2 are shown in Table 5.

Table 5. Accuracy Test Results

Mining data volume (GB)	Accuracy (%)		
	Design method	Method 1	Method 2
50	92.30	86.20	82.24
55	92.57	86.31	82.30
55	92.78	86.40	82.54
60	92.95	86.65	82.78
65	93.04	86.74	82.86
70	93.21	86.88	82.92
75	93.47	86.96	83.06
80	93.63	87.08	83.20
85	93.75	87.25	83.37
90	93.95	87.45	83.48
95	94.07	87.50	83.62
100	94.24	87.61	83.85

According to the test results in the table above, the accuracy rate of the design method increases with the increase of the amount of mining data, and its accuracy rate is higher than the test results of method 1 and method 2 in the same period. The main reason is that the method proposed in this article is based on multitasking learning to design a Chinese word segmentation sequence annotation method, which implements preprocessing of online teaching data for employment guidance courses, thereby improving the generalization ability and effectiveness of the model.

3.3.2 Matthews Correlation Coefficient Test Results

The Mathews correlation coefficient test results of the three methods are shown in Fig. 3.

According to the test results in Fig. 3, the Mathews correlation coefficient of the design method is significantly higher than that of Method 1 and Method 2, indicating that its prediction performance is better. The main reason is that the design method uses an improved K-means algorithm to mine the learning behavior information of students in the data after Chinese word segmentation, thereby discovering potential patterns and patterns hidden in the data.

Select 10 freshmen and 10 sophomores each, extract their behaviors related to online teaching in employment guidance courses, and analyze their comprehensive scores of employment guidance courses after using online teaching methods. The results are shown in Table 6.

According to Table 6, the difference between the predicted scores of the method applied in this article and the corresponding students' historical scores is less than 1.5 points. It has been proven that the online teaching and learning effect prediction method for employment guidance courses based on multitasking learning adds multiple types of

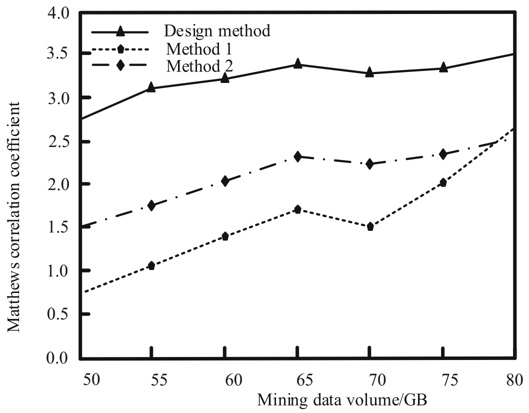

Fig. 3. Test Results of Matthews Correlation Coefficient

Table 6. Comprehensive Scoring of Employment Guidance Courses

Freshman	Historical Score	Predicted Score	Sophomore	Historical Score	Predicted Score
1	91.5	92.0	1	93.0	94.0
2	96.5	97.0	2	94.0	95.0
3	80.5	81.0	3	97.5	97.0
4	79.0	80.0	4	90.5	90.0
5	69.0	70.0	5	84.5	84.0
6	70.0	71	6	83.5	83.0
7	81.0	82.0	7	75.5	75.0
8	63.5	64.0	8	79.0	80.0
9	71.5	72.0	9	69.0	70.0
10	61.5	62.0	10	64.0	65.0

feature information to the EduHawkes model for learning effect prediction, achieving the prediction of online teaching and learning effect for employment guidance courses, improving the accuracy and interpretability of learning effect prediction, and having good prediction effects.

4 Conclusion

This paper takes the representation of learning behavior as the main line of the research, and takes the prediction of learning effect as the main research goal. It designs a prediction method of online teaching learning effect of career guidance courses based on multi task learning. The design method demonstrated good performance in predicting learning outcomes. The experimental results show that the accuracy of the design method gradually improves with the increase of mining data volume, and the overall accuracy is relatively high. In addition, the Matthews correlation coefficient of the design method is also significantly high, indicating that the model can accurately capture the correlation between students' learning behavior patterns and features. At the same time, the difference between the predicted score after the application of this method and the corresponding student's historical score is less than 1.5 points, which further verifies the accuracy and stability of the model.

However, despite the excellent performance of design methods in predicting learning outcomes, there are still some areas that need improvement. Future research should focus on addressing the following issues: how to better utilize multiple types of feature information to further improve prediction performance; How to combine knowledge and technology from other fields to further optimize prediction results. By addressing these issues, the accuracy and practicality of predicting the learning outcomes of online teaching in employment guidance courses can be further improved.

References

1. Zhang, N.: A brief analysis of how online teaching can develop steadily and be far-reaching in the new era-a case study of practical course for college music programs. Psychol. Res. **12**(7), 525–529 (2022)
2. Liu, P., Wang, X., Teng, F.: Online teaching quality evaluation based on multi-granularity probabilistic linguistic term sets. J. Intell. Fuzzy Syst. **40**(2), 1–20 (2021)
3. Bulut, T., Del, T.: Technology acceptance model: the perceptions of Turkish and Kazakh English teachers in online teaching. Int. J. Lang. Acad. **37**(37), 305–315 (2021)
4. Fitriani, S.S., Weda, S., Samad, I.A., et al.: Genre-based visualization through an online teaching platform: a strategy to engage with academic texts during the Covid-19 outbreak. XLinguae **14**(1), 270–288 (2021)
5. Prabha, M.S., Dinesh, K.V.: Distance learning and community of inquiry: how webinars for postgraduates are different from undergraduate online teaching sessions? CHRISMED J. Health Res. **7**(4), 287–288 (2021)
6. Panhwar, M., Parveen, S., Baig, Q.A., et al.: Educational challenges and their influence on online teaching methods in Pakistani students during COVID 19 era. Rawal Med. J. **46**(3), 677–680 (2021)
7. Jiang, T., Cui, L., Liu, et al.: The method of ship track prediction based on cluster analysis and Att-Bi-LSTM. Comp. Integr. Manuf. Syst. 39(8), 1–5+322 (2022)
8. Khanal, P.: Lived experience of online teaching during the COVID-19 pandemic: implications for curriculum and teaching. Interdisc. Res. Educ. **5**(1–2), 89–102 (2021)
9. Lou, Z.H., Wang, H.L., Sun, C.Y.: Simulation of academic development analysis model under background of education big data. Comp. Simul. **39**(12), 306–310 (2022)
10. Sato, E., Chen, J.C.: Rise to the occasion: the trajectory of a novice Japanese teacher's first online teaching through action research. Lang. Teach. Res. **25**(2), 306–329 (2021)