Research on Resource Classification Method of Mobile Education Platform for Physics Theory Teaching

Han Xu and Qiong Chen(✉)

Department of Basic Courses, Wuhan Donghu University, Wuhan 430212, China
MiuMiu1993@163.com

Abstract. There are various types of resources on mobile education platforms, but these resources often appear fragmented and lack systematicity and integration. Therefore, the classification research of physics theory teaching resources can help students better find and utilize resources that are suitable for their learning needs. A mobile education platform resource classification method for physics theory teaching is proposed to solve the problems of low accuracy and recall rate, as well as long resource classification time in traditional mobile education platform resource classification methods. Build a data collection architecture for physics theory teaching resources on the mobile education platform by using web page parsing module, text processing module, search strategy module, and supplementary mechanism module. Extract relevant resource features based on the collected data of physics theory teaching resources. Utilizing cost sensitive learning to improve Ada Boost ensemble learning algorithm, and combining resource feature extraction results to achieve resource classification on mobile education platforms. The experimental results show that the average accuracy of this method is 96.9%, the average recall rate is 96.8%, and the minimum time required for resource classification on mobile education platforms is 2.1 s. The classification results are reliable.

Keywords: Teaching physics theory · Mobile education platform · Resource classification · Cost sensitive learning · Ada Boost ensemble learning algorithm

1 Introduction

With the popularity of mobile Internet and the wide application of smart terminals such as smart phones, mobile education platform has become one of the hot spots in the current education field, which is particularly important for physics theory teaching [1]. Compared with traditional education, mobile education platform is not limited by time and space, has the advantages of flexibility and interaction, provides students with more learning possibilities, and can realize personalized teaching and the sharing of high-quality teaching resources. Mobile education platform resource classification refers to the classification of various learning resources on the mobile education platform in

G. Gui et al. (Eds.): eLEOT 2023, LNICST 543, pp. 283–299, 2024.
https://doi.org/10.1007/978-3-031-51465-4_19

accordance with a certain way, so that students can quickly find the required learning resources on the platform. These classification methods can be divided according to different dimensions, such as subject classification, knowledge point classification, difficulty level classification, non-curriculum resources classification, etc.The significance of studying the classification of mobile education platform resources is that through classification and organization, it can better improve students' learning efficiency and quality. The number of learning resources on mobile education platforms is huge [2]. If not classified and organized, students may feel confused and anxious during the process of searching and selecting learning resources, which affects learning efficiency and quality. Meanwhile, studying the classification of mobile education platform resources can also help improve the user experience of the platform, enhance students' learning experience, and enhance the brand value of the education platform. However, due to the numerous and complex knowledge points of physics theory, as well as the wide variety of educational resources, educators need to classify and manage resources to meet students' learning needs and achieve the best teaching results.

Aiming at the problem of resource classification on mobile education platform, literature [3] combined with a novel data enhancement method and calculated its attention matrix through attention mechanism according to the contribution of word vector to classification results, aiming at the characteristics of teaching material resource data set, such as rich text information, not obvious feature presentation and uneven sample distribution. Then the word vector matrix is combined with input into the model, so as to propose a text classification model Io META combined with attention mechanism, and use Io META to carry out deep learning of textbook resources, so as to realize resource classification. Reference [4] proposes an information resource classification method based on association rules to address the issue of current methods being unable to accurately analyze information feature relationships, resulting in low accuracy of information resource classification results. Utilize information gain to extract features of the information to be classified, and establish an evaluation function based on the difference in information gain. Using the distance from the feature to the corresponding category center as an association rule, deep mining of the internal relationship between information is achieved. By determining the itemset and using training, the classification of information resources is completed.

However, problems such as low accuracy of classification results, low recall rate and long classification time are found in the application of these two methods, leading to poor classification effect of mobile education platform resources. To solve this problem, this paper proposes a resource classification method for mobile education platform oriented to physics theory teaching. Through reasonable classification and integration of resources, it provides educators with effective educational resource management schemes, which aims to ensure effective management and utilization of educational resources and bring more and better learning resources to students. At the same time, the method will also consider students' academic needs and interests, in order to better meet the learning needs of different students. The structure of this article is as follows:

1) Build a data collection architecture for physics theory teaching resources on the mobile education platform using web page parsing module, text processing module, search strategy module, and supplementary mechanism module, and obtain relevant data collection results.
2) Extract relevant resource features based on the collected data of physics theory teaching resources. Cost sensitive learning is used to improve Ada Boost Ensemble learning algorithm, and resource feature extraction results are combined to achieve resource classification of mobile education platform.
3) The accuracy and recall rate of resource classification on mobile education platforms, as well as the time spent on resource classification, were used as indicators to comprehensively test the effectiveness of this method.
4) Summarize the entire article and draw a conclusion.

The main contributions of resource classification research on mobile education platforms for physics theory teaching are as follows:

1) Personalized resource recommendation: This study can achieve personalized recommendation of physics theory teaching resources by analyzing students' learning behavior, interest preferences, and learning needs. By associating student portraits with resource classification, students' personalized learning needs can be better met, and learning effectiveness and motivation can be improved.
2) Resource integration and organization: In response to the diversity and fragmentation of physics theory teaching resources, this study can integrate and organize different forms of resources to build a more complete and coherent learning resource chain. By categorizing and organizing resources such as videos, PPTs, PDFs, and codes, we aim to provide students with a more diverse learning experience and support.
3) Resource quality assessment and assurance: This study can establish standards and indicator systems for resource evaluation, evaluate and screen resources through automated or semi-automatic methods, and improve learners' trust and effectiveness in resource utilization. By evaluating and ensuring the quality of resources, high-quality physics theory teaching resources can be provided to improve learning outcomes.
4) Teaching strategy optimization: By analyzing the behavior data of learners on mobile education platforms, it is possible to understand the usage and effectiveness of different resources by students. Based on these data, teaching strategies can be optimized, resource classification and recommendation strategies can be adjusted, and teaching effectiveness and student satisfaction can be improved.
5) Cross platform sharing and interoperability: This study can explore the establishment of unified resource classification standards and norms, achieving resource sharing and interoperability among different mobile education platforms. This helps to promote the cross platform circulation and sharing of resources, improve their accessibility and sustainability.

2 Design of Resource Classification Method for Mobile Education Platform

2.1 Data Collection Architecture of Physics Theory Teaching Resources on Mobile Education Platform

The data collection architecture for physics theory teaching resources on the mobile education platform designed in this article adopts a B/S architecture as a whole. The MySQL database stores crawling data, the Redis cache database serves as a middleware database, the Neo4j graph database stores Fin Graph knowledge graphs, and the front-end page developed based on the Vue.js framework is used to complete user interaction. The back-end server provides logical processing, and the database stores the collected webpage information, The functions are mainly divided into web page parsing module, text processing module, search strategy module, and supplementary mechanism module. The data collection architecture is shown in Fig. 1.

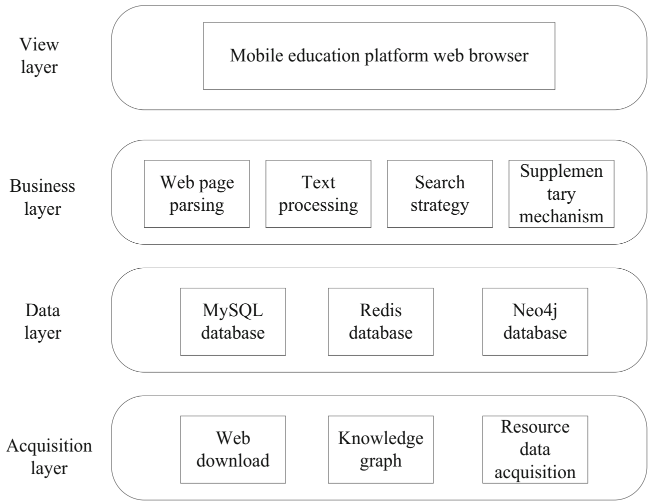

Fig. 1. Data collection architecture of physics theory teaching resources on mobile education platform

The user interacts with the view through the web browser. The browser is responsible for the front page display, and the back end is responsible for responding to the corresponding operations of the front end. The business layer includes webpage parsing module, text processing module, search strategy module and supplementary mechanism module. The data layer includes My SQL database, Redis database and Neo4j database. My SQL and Neo4j are mainly used to store subject-related webpage text data and Fin Graph knowledge graph, while Redis database is used for url reprocessing.The collection layer includes data collection function, which downloads webpage content based on URLs. Fin Graph assists in extracting key phrases from webpage text for topic semantic similarity calculation. According to the supplementary mechanism of the business layer, high correlation webpage crawls are extracted from triples to update Fin Graph, thereby building relevant knowledge graphs and achieving data collection of physics theory teaching resources on mobile education platforms.

(1) Design of web page parsing module

The webpage parsing module is responsible for extracting text, page title and link information from the obtained HTML source code [5], and removing redundant data and label structure of physical theory teaching resources on mobile education platform. This paper uses Xpath selector for preliminary screening. Xpath selector uses path expressions to select nodes or node sets in webpage documents, and extracts field data to webpage documents of mobile education platform by locating target nodes. Compared to CSS selectors, Xpath selectors' expressions are more concise, flexible, and efficient in finding complex elements. After the preliminary screening, regular expressions are used to extract the specific requirements.

(2) Text processing module

The text processing module is responsible for processing the extracted text of mobile education platform web pages. As the extracted text may contain residual web page tags, irrelevant punctuation mark and characters, and stop words, the text of the web page text should be cleaned to get clean physical theory teaching resource data of mobile education platform by removing miscellaneous content. Afterwards, the APKGram algorithm and Fin Graph knowledge graph were combined to extract key phrases from the cleaned web page text. Finally, word vector training is conducted based on relevant corpora to extract key phrase segmentation from the webpage body and map the segmentation to the training results to obtain the word vector for each word. If there are missing values, a word vector representation with the same dimension is randomly initialized. The word vectors of each key phrase are the average of the constituent word vectors, and the set of word vectors of these key phrases serves as the text feature vectors of the entire mobile education platform webpage.

(3) Search strategy module design

The search strategy module has the function of judging strategy based on the mixed content and link structure of mobile education platform. New links found by theme web crawler in the process of crawling web pages can be divided into two categories. One is directory page link, which is usually seen as the next page link. The other is the web content link, that is, the link that needs to parse the extracted fields of the web page, pointing to the actual mobile education platform physical theory teaching resource data information page. Theme web crawler [6, 7] extracts multiple content page links and web contents from each directory page link, selects the required links according to the set link structure rule expression and adds them to the cache database Redis to be downloaded, and removes the interference of irrelevant pages. And by calculating the semantic similarity between the content of the web page and the topic to decide whether to store the text of the web page in My SQL database.

This article uses the Redis database to store the URL queue. The Redis database is a high-performance non relational database that can support multiple data structures and can support persistence logs based on memory. The key of the crawler from

Redis is next_ When the link extracts its value value, the initial link starts running. The downloader returns a response to obtain the webpage content text. During the parsing of the webpage content, the following judgment process is performed:

1) Determine whether the webpage content has a link to the next page, and if so, store it in the Redis download queue;
2) Judge whether the webpage content matches the detail content page. If so, consider the page link and text content. For the link, use the set regular expression to extract the page link that meets the requirements and store it in the Redis download queue; For the text content, the topic similarity calculation based on the mobile education platform physics theory teaching resources text content is used to determine whether it is stored in the database. By comparing the semantic similarity between the text feature vector obtained from the text processing module and the topic feature vector, the semantic matching degree between the text content and the topic is calculated. The cosine similarity is used to calculate the topic similarity of the webpage text. If the calculated topic similarity is greater than or equal to the set threshold, the text content of the web page and its related information will be stored in My SQL database, otherwise it will not be recorded. After completing the above steps, if the next_link list is not empty, the crawler continues to fetch urls from the Redis queue for the above judgment steps until the waiting queue is empty. If the next_link list is empty, the process ends.

(4) Supplementary mechanism module

The supplementary mechanism module further supplements the Fin Graph knowledge graph based on a certain amount of physical theory teaching resource text data collected by crawlers, making the knowledge graph more complete. This article sets up to crawl 1000 physics theory teaching web pages, extract entity and relationship triplets through a joint extraction learning model of constructing a knowledge graph, and then perform entity disambiguation and co referential resolution to complete knowledge fusion. The fused "entity relationship entity" triplets are obtained and stored in the Neo4j graph database. Its significance lies in improving the accuracy of extracting key phrases from webpage text based on the supplemented Fin Graph knowledge graph, thereby affecting the crawler's ability to crawl webpage content that is more in line with the theme. The two complement each other [8].

(5) My SQL database is mainly used to store the webpage information of mobile education platform collected by crawler, including webpage URL, title, text, publication time, source and storage time.

The main detailed process of themed web crawler for data collection of physics theory teaching resources on mobile education platform is described as follows:

1) The crawler reads the initial URL in the URL queue and sends the request to the server. After obtaining the response returned by the server, the webpage is downloaded and the HTML source code is obtained [9];
2) The webpage parsing module analyzes the obtained HTML source code, removes irrelevant content and structure, and retains the webpage text;
3) The text processing module processes the text of the web page to some extent, and extracts key phrases in the text with the help of Fin Graph to get the feature vector of the text of the web page;

4) The search strategy module is used to determine a strategy based on a mixture of web page content and link structure. It calculates the semantic similarity between the web page text feature vector and the input topic vector. If the set threshold is met, the relevant information of the text is stored in the database, and the page links that meet the requirements are extracted and stored in the URL queue according to the rules.
5) Finally, determine whether the stop condition has been met. If the URL queue is empty, the program ends. If the URL queue is not empty, the next URL is passed to the webpage download section and the program continues to run.
6) The supplementary mechanism module stores a certain amount of physical theory teaching resource text data on a mobile education platform in a database, and then inputs these text data into a joint extraction learning model to obtain triplets, in order to supplement the Fin Graph knowledge graph and achieve data collection of physical theory teaching resources on mobile education platforms.

2.2 Resource Preprocessing of Mobile Education Platform

Based on the collected data, extract the characteristics of physical theory teaching resources on mobile education platforms to ensure subsequent classification accuracy.

The TF-IDF algorithm was proposed by Professor Salton in 1973 [10]. It is a statistical method used to evaluate the importance of a word or phrase to a document or category in a corpus. The main idea is that if a word or phrase appears frequently in one category and rarely in other categories, it is considered to have good category differentiation ability and is suitable for resource feature extraction.

TF represents the total frequency of a feature item t appearing in document d. In the early stages of artificial intelligence, the *TF* algorithm is usually chosen for text processing of physical theory teaching resources on mobile education platforms. The formula is as follows:

$$TF(t, d) = \frac{f(t, d)}{\sum_{k=1}^{n} f(w_k, d)} \tag{1}$$

where, $f(t, d)$ represents the total number of occurrences of feature item t in document d, and $f(w_k, d)$ represents the total number of occurrences of feature item d.

IDF is used to see if a feature item t is ubiquitous in the document. The main idea is: if the fewer documents contain the feature item t, the larger the IDF value, indicating that the feature has good distinguishing ability. The formula is as follows:

$$IDF(w) = \log\left(\frac{N}{1 + df_t}\right) \tag{2}$$

Among them, N represents the total number of documents in the corpus, and df_t represents the number of documents containing feature item t.

TF-IDF evaluates feature terms from two aspects during calculation. The weight of feature item t increases proportionally with its frequency in the corpus, but at the same time decreases inversely with its frequency in the corpus. The complete formula is as follows:

$$IF - IDF = IF(t, d) \times IDF(t) \tag{3}$$

It can be seen from the definition of IDF formula that IDF mainly considers the category differentiation of feature item t from the perspective of the whole corpus set, and lacks consideration of the distribution of feature item t among classes, thus affecting the classification accuracy.

Since the traditional IDF calculation method does not consider the distribution between classes, an improved IDF calculation method is proposed here. The logarithmic molecule of the IDF formula is changed from the number of the whole document set to the number of all documents under a certain category, and then the IDF value under each category is calculated separately, and then the IDF value of all categories is used for variance calculation $D(t)$, the specific formula is as follows:

$$D(t) = \frac{\sum_{i=1}^{m} (IDF(t, c_i) - IDF(t))^2}{m} \tag{4}$$

Among them, $D(t)$ represents the distribution of the concentration degree of feature item t in different categories of text sets. It can be seen that the value of $D(t)$ is inversely proportional to the distribution of feature item t in different categories. The less concentrated the feature item is, the better the degree of differentiation, and the more representative it is of a class. The improved formula is as follows:

$$IF - IDF = IF(t, d) \times IDF(t) \tag{5}$$

The whole mobile education platform physics theory teaching resource data set is trained by the word vector model, so as to obtain the word vector of each feature item t. The word vectors of feature item t appearing in text w_i are summed respectively to obtain the sentence vector $R(w_i)$ of text w_i. Where, $word2vec(t)$ indicates the word vector of the feature item t.

$$R(w_i) = \sum_{t} word2vec(t), t \in w_i \tag{6}$$

Then, the improved TF IDF ‘formula is used to weight the word vector and obtain $weight\_R(w_i)$.

$$weight\_R(w_i) = word2vec(t) \times weight_t, weight_t = TF - IDF \tag{7}$$

The feature extraction process of physics theory teaching resources on mobile education platform is as follows.

Input: mobile education platform physics theory teaching resource data set c, document set D_j of each category, and each text d_i, optimal feature item set T.

Output: Resource feature extraction results.

1) The sentences in the data set of physics theory teaching resources on the mobile education platform are processed by word segmentation, words that do not belong to the optimal feature set T are filtered out, only words belonging to the optimal feature set T are reserved, and the results are saved in the Word2Vec training model;
2) Train the Word2Vec model on the model and set the size length to 50;
3) Calculate the IDF values of each feature item in the optimal feature set in different categories of documents;
4) Perform variance calculation on the IDF value of each feature item t, and then multiply it by its corresponding TF value to form the improved TF-IDF';
5) The word vectors trained using the Word2Vec model represent each text, and TF-IDF' is used to weight each feature item that appears in the text;
6) Return the text vector representation matrix and use it to extract the characteristics of physics theory teaching resources on mobile education platforms.

2.3 Resource Classification of Mobile Education Platform Based on Improved Integrated Learning

Ada Boost ensemble learning algorithm adopts Boosting ensemble idea, which links up multiple classifiers in series. With two weights, namely sample weight and base classifier weight, Ada Boost ensemble learning algorithm can achieve key learning of the misclassified samples, so as to improve the classification effect of resources on the mobile education platform. The specific process is as follows.

(1) Initialize the sample weights of the training set of resource data of the mobile education platform, so that all samples have the same weights

$$D_1 = \left(W_{1,1}, W_{1,i}, ..., W_{1,m}\right), W_{1,i} = \frac{1}{m}, i = 1, 2, ..., m \tag{8}$$

(2) Train the base classifier using a training set of mobile education platform resource data samples with a weight of D_t to obtain the t weak classifier $D_t(x)$.

(3) Calculate the classification error rate of $D_t(x)$ in the mobile education platform resource data training set:

$$e_t = P(G_t(x_t) \neq y_i) = \sum_{i=1}^{m} W_{ti} I(G_t(x_t)) \neq y_i \tag{9}$$

The weight of weak classifier $D_t(x)$ in strong classifier is calculated according to the classification accuracy of weak classifier

$$\alpha_t = \frac{1}{2} \ln \frac{1 - e_t}{e_t} \tag{10}$$

(4) Update the weights of the samples in the training dataset based on the results of the previous iteration

$$D_{t+1} = \left(W_{t+1,1}, ..., W_{t+1,i}, ..., W_{t+1,m}\right) \tag{11}$$

$$W_{t+1,i} = \frac{w_{t,i}}{Z_t} \exp(-\alpha_t y_i G_t(x_i)) \tag{12}$$

where, Z_t represents the gauge factor.

(5) Construct the linear combination of weak classifiers for resource classification of mobile education platform, specifically as follows:

$$f(x) = \sum_{i=1}^{T} \alpha_t G_t(x) \tag{13}$$

(6) The final strong classifier for resource classification on mobile education platforms is as follows:

$$G(x) = sign(f(x)) = sign\left(\sum_{i=1}^{T} \alpha_t G_t(x)\right) \tag{14}$$

It can be seen from the above process that the key of Ada Boost integrated learning algorithm is adaptive dynamic updating of sample weights, and samples of resource classification results on different mobile education platforms have different weight updates.

When the sample prediction is correct:

$$W_{t+1,i} = \frac{w_{t,i}}{Z_t} \exp(-\alpha_t) \tag{15}$$

When the sample prediction is incorrect:

$$W_{t+1,i} = \frac{w_{t,i}}{Z_t} \exp(\alpha_t) \tag{16}$$

where $\alpha_t \geq 0.5$, when the classification of resource data samples on the mobile education platform is correct, the weight of the samples will be reduced; when the classification of resource data samples on the mobile education platform is wrong, the weight will also increase, so as to realize the focus of learning on the wrong samples. However, in the unbalanced data classification problem, due to the small number of minority samples, the standard machine learning algorithm which takes the overall optimization as the optimization strategy is easy to classify the minority samples incorrectly. Therefore, through multiple learning, better identification results of minority samples will be obtained compared with other ensemble learning models.

The traditional machine learning classification algorithm takes the improvement of the overall classification accuracy as the optimization goal, and treats the misclassification cost of each type of sample equally in the loss function, but in the unbalanced classification problem, this is not very appropriate. Moreover, in imbalanced data, the error cost of misclassifying minority samples into majority samples is much higher than that of misclassifying majority samples into minority samples. For imbalanced classification problems, different types of misclassification costs should be treated differently based on the consideration of misclassification costs, in order to achieve the minimum comprehensive misclassification cost. Therefore, in the imbalanced data classification problem, the difference in misclassification costs of samples from different categories is fully considered, and cost sensitive learning is introduced to transform the learning objective of the classification algorithm from reducing overall error to reducing classification costs. Cost sensitive learning can be divided into input stage introducing cost sensitive learning, algorithm stage introducing cost sensitive learning, and output node introducing cost sensitive learning based on the timing considered.

The cost sensitive learning is introduced into the algorithm, the algorithm is improved, the loss function is modified, and the change of the optimization objective of the loss function is used to realize the effective identification of a few types of samples while ensuring the overall accuracy, and FP and FN in the confusion matrix are treated differently. That is:

$$R_L = R(i,j) * L(FP) + R(j,i) * L(FN) \quad (17)$$

Among them, $R(i,j)$ represents the loss cost when Class j is misclassified as Class i, $R(j,i)$ represents the loss cost when Class i is misclassified as Class j, and $R(i,j) \neq R(i,j)$. x represents an instance. If $j = i$, it indicates that the model classifies mobile

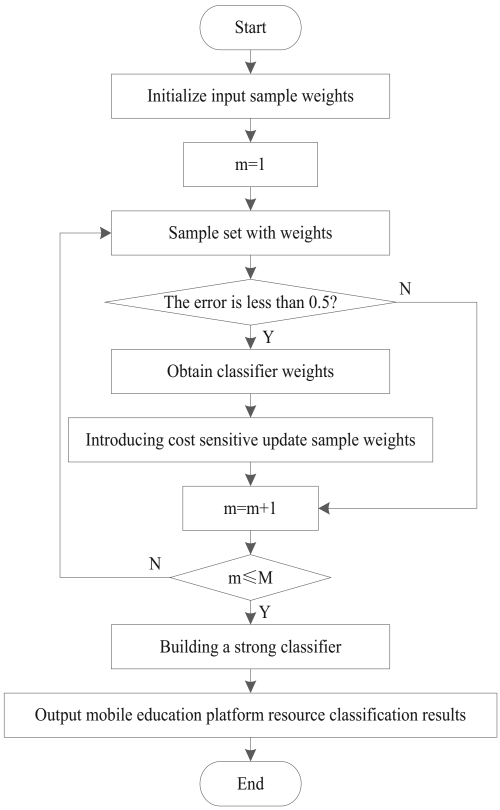

Fig. 2. Resource Classification Process for Mobile Education Platform

education platform resource samples correctly. The cost sensitive classification loss function involved in this concept no longer takes the overall mobile education platform resource classification accuracy as the optimal goal, but the expected cost of classification is the smallest, namely:

$$\arg\min E(x, y) \ : \ C\left[t_{yR(x)}\right] \tag{18}$$

where, E represents the expected cost of classification, C represents the unbalanced data distribution, $R(x)$ represents the structural risk value of x, and $t_{yR(x)}$ represents the misclassification cost of a certain category. In order to achieve the iteration process to achieve the specified class of samples in the loss function to increase the discourse power, to achieve the focus of the iteration process. The resource classification process of mobile education platform is shown in Fig. 2.

3 Experimental Design

The configuration of the experimental platform in this article is diverse, mainly conducting simulation experiments on computers. The specific operating system is Windows 10, and Pandas is used as a tool for reading and preprocessing file data. The specific experimental environment is shown in Table 1.

Table 1. Experimental environment

Environment name	Environment configuration
Operating system	Windows 10
CPU	Inter CoRE i5-5200U
GPU	NVIDIA GTX 1650
Memory	16 GB
Programming language	Python 3.8
Integrated development environment	Pych

The types of experimental sample data selected during this experiment are as follows:
Type 1: Text Data

1) Image data: Course name and introduction of physics theory: Describe the name and main content of the course to help students quickly understand the overview of the course.
2) Textbook content: including chapters, class hours, and knowledge points, to help students learn and master the course content.
3) Study notes: This includes notes, insights, and summaries of students' physics theory courses, which helps them review and consolidate course knowledge.
4) Teacher's handouts and lesson plans: including teacher's handouts and classroom lesson plans to assist teachers in lesson preparation and teaching.

5) Evaluation and feedback: This includes students' evaluation and feedback on physics theory courses, which helps to provide feedback, improve and improve course quality.

Type 2: Audio data

1) Course audio: including the recording and explanation of the physics theory course, students can listen to the class anytime and anywhere through mobile devices.
2) Teaching audio: including teachers' lectures and answers, to help students better understand and master the content of physics theory course in learning.
3) Evaluation and feedback: including students' evaluation and feedback on the physics theory course, which is helpful to provide feedback and improve and perfect the quality of the course.
4) Voice interaction: including voice assistant and voice recognition, to help students to carry out interactive learning, improve the interest and effect of learning.

Type 3: Video Data

1) Course video: including videos recorded for physics theory teaching, which can be watched anytime and anywhere through mobile devices.
2) Teaching video: including teacher's explanations, demonstrations, problem-solving, etc., to help students better understand and master the content of physics theory courses during learning.
3) Evaluation and feedback: This includes students' evaluation and feedback on physics theory courses, which helps to provide feedback, improve and improve course quality.
4) Virtual simulation videos: including virtual simulation experiments, virtual demonstrations, etc., can provide more vivid and vivid learning methods.

Type 4: Evaluation data

1) Evaluation data of mobile education platform refers to students' evaluation data on courses and teaching, which mainly includes the following contents:
2) Course evaluation: including the evaluation of the difficulty, learning content, curriculum setting and other aspects of the physics theory course.
3) Teaching evaluation: including the evaluation of teachers' explanation, solution, guidance, interaction and other aspects.
4) Teaching environment evaluation: including the evaluation of physical theory teaching classroom environment, learning equipment, learning atmosphere and other aspects.
5) Evaluation of learning experience: including evaluation of learning style, learning effect, learning resources, sense of experience and other aspects.

Type 5: Behavioral Data

The behavior data of mobile education platforms refers to the learning and operational behavior data of students on the platform, mainly including the following content:

1) Learning records: including students' learning duration, visit frequency, learning content, learning progress, etc.
2) Learning interaction: including interaction between students and teachers, interaction between students, etc., such as questioning, discussion, etc.

3) Learning outcomes: including homework, exam, experimental data, and other outcome data during the student learning process, reflecting the student's learning outcomes in a timely manner.
4) Operational behavior: including students' click behavior, search behavior, collection behavior, evaluation behavior, etc., reflecting students' interest and inclination towards learning resources.

Each experimental sample data category contains 10,000 records, with 80% randomly designated as the training set and 20% as the test set respectively. After training the simulation software, the data of the test set is input into the simulation software to obtain the relevant experimental results. The method of literature [3], literature [4] and the method in this paper are used as experimental comparison methods, and the accuracy rate, recall rate and classification time of different methods are compared to verify the effectiveness of this method.

Classification accuracy rate refers to the proportion of all samples classified as a certain category, which really belongs to that category. It is one of the important indicators to evaluate the performance. The higher the classification accuracy rate is, the higher the proportion of the samples of this category is correctly classified by the classification method, and the higher the classification accuracy is. The comparison results of the accuracy rates of the methods in literature [3], literature [4] and the methods in this paper are shown in Table 2.

Table 2. Precision Comparison Results (Unit:%)

Number of experiments	Reference [3] Method	Reference [4] Method	Proposed method
1	75.6	81.3	96.8
21	76.3	80.6	97.2
41	74.9	84.7	98.7
61	72.5	81.3	95.6
81	73.4	81.4	96.1
Mean Value	74.5	81.9	96.9

By analyzing the results in Table 2, we can see that with the increase of experiment times, the accuracy rates of the three methods all show a fluctuating trend. Among them, the average accuracy rate of the method in literature [3] is 74.5%, the average accuracy rate of the method in literature [4] is 81.9%, and the average accuracy rate of the method in this paper is 96.9%, which is much higher than that of the experimental comparison method, indicating that the method has a high accuracy for the classification of resources on mobile education platform and good practical application effect. The reason is that this method uses cost sensitive learning to improve Ada Boost Ensemble learning algorithm to achieve resource classification of mobile education platform by extracting relevant resource characteristics. Therefore, the classification accuracy of this method is high.

Classification recall rate refers to the percentage of all samples that actually belong to a class that are classified into that class. It is another important index to evaluate the performance of the classification method. The higher the classification recall rate is, the better the classification method can capture the samples of this category, and the better the classification effect is. The comparison results of the recall rate of the method in literature [3], literature [4] and the method in this paper are shown in Table 3.

Table 3. Comparison results of recall rate (unit: %)

Number of experiments	Reference [3] Method	Reference [4] Method	Proposed method
1	86.3	79.6	98.7
21	85.4	86.3	95.6
41	85.6	85.2	94.7
61	84.8	81.3	97.6
81	81.7	80.7	97.5
Mean Value	84.8	82.6	96.8

By analyzing the results in Table 3, we can see that with the increase of experiment times, the recall rates of the three methods all show a fluctuating trend. Among them, the average recall rate of the method in literature [3] is 84.8%, that of the method in literature [4] is 82.6%, and that of the method in this paper is 96.8%, which is much higher than the experimental comparison method, indicating that the method has a higher recall rate and better classification quality for the mobile education platform resources. The reason is that this method uses cost sensitive learning to improve Ada Boost Ensemble learning algorithm, and combines the resource feature extraction results to achieve resource classification of mobile education platform. The improved Ada Boost Ensemble learning algorithm has the advantages of improving classification accuracy, handling unbalanced data sets, enhancing robustness, maintaining interpretability and adjustability, and using parallel computing capabilities in resource classification. These advantages make the improved Ada Boost algorithm an effective and feasible method for resource classification tasks, resulting in a higher recall rate of classification results.

Classification time index refers to the time index required to classify resources of mobile education platform in the process of data processing. In order to improve the efficiency and accuracy of classification, these indexes should be evaluated and optimized during resource classification of mobile education platform. The comparison results of resource classification time of mobile education platform based on literature [3], literature [4] and this paper are shown in Fig. 3.

The analysis of experimental results in Fig. 3 shows that with the increase of experiment times, the resource classification time of the three methods of mobile education platform also changes to some extent. Among them, the minimum time of resource classification of mobile education platform in literature [3] is 5.3 s, the minimum time of resource classification of mobile education platform in literature [4] is 4.5 s, and the

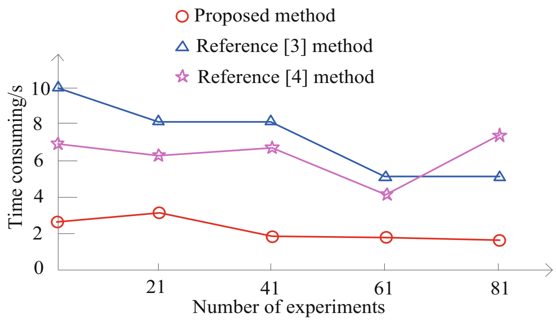

Fig. 3. Comparison results of classification time consumption

minimum time of resource classification of mobile education platform in this paper is 2.1 s, indicating that the resource classification of mobile education platform in this paper is shorter and more efficient. The main reason is the introduction of the improved Ada Boost Ensemble learning algorithm, which can realize parallel computing in resource classification, so the classification time of this method is always kept at a low level.

4 Conclusion

With the popularization of mobile internet and intelligent terminals, more and more educators are beginning to realize the importance of mobile education platforms in education and teaching. Especially in the field of physics theory teaching, the advantages of mobile education platforms are particularly evident. Through mobile education platforms, students can learn at any time, anywhere, and according to their own learning progress and interests. The research on resource classification methods for mobile education platforms for physics theory teaching has become one of the hotspots in the current education field. Therefore, this article proposes a mobile education platform resource classification method for physics theory teaching. The experimental results show that the accuracy and recall of this method are high, and the classification time is short. It can achieve accurate and fast classification of mobile education platform resources, promoting further improvement of physics theory teaching level. However, the resource classification methods studied for mobile education platforms still face some challenges and shortcomings, such as a lack of consistency and standardization, improvement in classification accuracy, and difficulty in resource evaluation and quality assurance. Future research can focus on addressing these issues, improving the accuracy, personalization, and quality of resource classification, in order to provide a better learning experience and support.

Acknowledgement. 1. 2022 Hubei Provincial Department of Education Science and Technology Research Program Project B2022315.

2. 2021 Wuhan Donghu University Youth Fund Project Approval (Natural Science) 2021 dhzk0077.

References

1. Lin, H.: The construction of mobile education in cloud computing. Procedia Comput. Sci. **183**(8), 14–17 (2021)
2. Wu, Y., Chen, J.: Realization of mobile education resource sharing method based on wireless broadband connection. Sci Program. **7**(1), 1–10 (2021)
3. Guo, S.W., Chen, J.H.: Textbook classification method of index of moral education based on deep learning. Comput. Moder. **9**, 106–112 (2021)
4. Zhou, M.: A Classification method for personnel archive information resources based on association rules. Jiangxi Commun. Sci. Technol. **3**, 28–30 (2021)
5. Liu, D.L., Lv, M.: User information extraction based on web crawler and associated big data. Comput. Simul. **38**(8), 482–486 (2021)
6. Yu, L., Li, Y., Zeng, Q.: Design of topic Web crawler based on improved PageRank algorithm. J. Phys: Conf. Ser. **1754**(1), 012210–012217 (2021)
7. Kuze, D.: Classification of diversified web crawler accesses inspired by biological adaptation. Int. J. Bio-Insp. Comput. **17**(3), 1–12 (2021)
8. Luo, M., Lin, J.: Research on the impact of credit, brand and service recovery on online sales based on web crawler technology and regression analysis methods. J. Phys: Conf. Ser. **1955**(1), 012083–012095 (2021)
9. Han, D.H., Lee, Y.K.: Design of action-based web crawler structural configuration for multi-website management. KIISE Trans. Comput. Pract. **27**(2), 98–103 (2021)
10. Yu, H., Ji, Y., Li, Q.: Student sentiment classification model based on GRU neural network and TF-IDF algorithm. J. Intell. Fuzzy Syst. **40**(2), 2301–2311 (2021)