Research Article

Recommendation of English Reading in Vocational Colleges Using Linear Regression Training Model

Haibo Xie

International School, Ningbo City College of Vocational Technology, China

Correspondence should be addressed to Haibo Xie; xiehaibo@nbcc.edu.cn

Received 23 March 2022; Revised 12 April 2022; Accepted 22 April 2022; Published 12 May 2022

Academic Editor: Mian Ahmad Jan

Copyright © 2022 Haibo Xie. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

High-quality vocational education is a vital basis for China’s efforts to develop high-quality teachers and technical and skilled workers. The increasing development of contemporary vocational education is given higher importance in the new era. We proposed the stick to moral education while carefully integrating it with the requirements of changes in technology and modernization, continuous improvements in the quality of college vocational education, and the development of more conceptual and skilled talented individuals with both political integrity and the capacity to build modernization. Using artificial intelligence technology to challenge these difficulties in the context of social informatization can play an important role in enhancing the teaching quality of high vocational education. With the advent of the Internet era, students and parents have an increasingly strong demand for online independent vocational English reading materials, although there is some English learning software that provides online vocational English reading materials in the education market. However, most of this software only changes the reading form of paper vocational English reading materials but fails to completely meet the real demand of students for vocational English reading materials, so how to carry out personalized recommendation reading according to the real reading situation and reading preferences of students is a vital problem to be resolved. In this background, this paper studies the personalized recommendation technology, in the use of recommendation algorithms on the basis of association rules to obtain higher vocational English reading content recommended the strongest association rules, on the basis of combining user-based collaborative clarifying recommendation algorithm. The linear regression model combined with the users reading interest and personalized recommendation of higher vocational English reading process is optimized, improving the accuracy of English reading recommendation in vocational colleges.

1. Introduction

Technical and vocational colleges’ prominence as primary scorers in chain experts has expanded dramatically. For the purposes of enhancement, the development has done different improvements. This may be observed in the government’s efforts, which range from the purchase and enhancement of equipment to the quality of education and teachers and the modification of the program [1]. Therefore, college education organizations that prepare instructors must build strategies to address the difficulty of implementing English education in colleges. Vocational education is an essential component of the national education scheme and the development of human resources. Higher vocational college movements should embrace possibilities, maintain education, promote teaching method modification and invention, improve teaching quality, and produce more increased, professionally qualified talent [2]. In the era of social information, artificial intelligence techniques for evaluating the quality of education in academic and vocational colleges have become an increasingly significant mechanism for higher vocational colleges to enhance the quality of education and talent development [3]. In the last several years, with the rapid development of the Internet information phase, traditional education is gradually transforming into Internet education [4]. As to strongly promote the development of higher vocational quality education in English, one of the important ways of English reading teaching, the purpose is the use of interaction between text and graphics to express the connotation of the story, to describe their
emotions and context, thus causing the student to achieve emotional resonance; this got the attention of more and more students and education institutions [5]. However, with the increasing demand for online English reading materials, a significant problem has gradually emerged; that is, students are unable to cultivate their attention in English and improve their education ability in the light of their reading situation and interest changes [6, 7]. Some components of pronunciation are overlooked by Chinese English learners, and there is no shortage of qualities in spoken English that are ignored, making it the bottleneck for increasing their spoken English level. In a typical aptitude test, it is unusual to record the students’ responses now [8, 9]. It is impossible to recreate each examinee’s assessment environment for other instructors or students at the time, much along with advice on the examinee’s answers, or other applicants cannot profit from the examinee’s experiences and learning [10].

This study will examine the meaning of AI technology and its natural language comprehension concepts, as well as the benefits of AI in the area of education. Furthermore, it suggests how artificial intelligence technologies can be used in English language learning in the input and output stages [11]. The findings revealed that academics had a largely positive desire to employ reversed teaching, with individual differences based on gender, the reputation of colleges, involvement, and the kind of English courses [12]. Their motivation to use flipping teaching was also predicted by their satisfaction with their objectives, according to with data. When teachers had high levels of self and support, identified regulation was significantly more strongly linked to use purpose, but external regulation was not related to use intention when teachers had greater levels of flipped teaching beliefs [13]. The whole planning and implementation of an intelligent spoken English correction system are described in this work, with a focus on the techniques for performing the functions from spoken English assessment, question paper management, and marking. A language emotion identification approach based on a multifeature fusion of sample entropy (SE) and Cepstrum coefficients (MFCC) is suggested. The support vector machine is used to analyze SE and its data, and also MFCC, and determines the likelihood that they correspond to one of six emotions. According to the experimental assessment, the English recognition system reported in this research has apparent performance increases in different measures [14].

In practical projects, how accurately discovering students’ interest preferences in English reading materials and making personalized English reading materials recommendations for students is the key to the whole English reading materials recommendation system [15, 16]. Students are in an age of active thinking and changeable interests, especially vocational students, who have a great degree of cognition of their interests and learning ability [17, 18]. In this context, this paper studies the theory and scheme of combining a personalized recommendation algorithm and linear regression training model with students’ reading interests, aiming to improve the recommendation effect of the vocational English reading recommendation system.

The following are some of the research contributions of this paper.

(i) First of all, we suggest using a multilinear regression technique at a vocational college for English reading recommendations

(ii) We do comprehensive testing on extremely popular and vocational institutions with different numbers of users, comparing our method to different algorithms for reading materials in vocational English

(iii) The proposed model uses expounds on the text classification of English reading materials in the vocational college prediction model which is better than the other conventional models

(iv) Using the traditional association rule-based and user-based collaborative filtering recommendation method, the system’s recommendation scheme based on the linear regression training model is improved

The rest of this paper is arranged in a logical order: Section 2 represents the related work, Section 3 shows the linear regression training model and recommendation framework, Section 4 represents the analysis of English reading material, Section 5 represents the high vocational English reading material accurate recommendation, and Section 6 captures the simulation evaluation. Finally, Section 7 takes the research work to a conclusion.

2. Relative Work

Using a grouping of artificial intelligence and the Internet, this paper explored the teaching of varied education modes based on WeChat and SPOC classrooms on the Internet. Based on the findings of the investigation, we suggest a new varied English teaching mode based on such stages to highlight the combination of artificial intelligence and the Internet. In the second principle, future English instructors must continue to develop their talents in the integration of educational resources, creative educational integration, and the development of learning activities. During college English classroom instruction, professional-quality enhanced performance and identity of personality attraction and actively responded to the era of AI [19]. This study will investigate the significance of AI technology and its natural language comprehension techniques and the benefits of AI in the area of education. Furthermore, it proposes concrete steps for incorporating AI technologies into English language learning in the input and output stages [20]. According to the study, instructors and other educational performers should help learners to be using deep and consequential learning techniques rather than the demotivating surface approach for them to achieve better learning outcomes, both quantitatively and qualitatively, and to acquire competitive qualifications for their future career goals [21]. This research presented a data mining technique to improve the satisfaction of students with high vocational education teaching quality. We created
a survey to measure the satisfaction of students with the fundamental entrepreneurship curriculum’s teaching quality. We utilize investigation data from vocation education as an instance to analyze the present state of the teaching quality of the fundamental entrepreneurship curriculum using mining technology analysis tools. The findings of this study may be utilized in high vocation college education for student information, educational approach, student satisfaction with overall education, and educational quality, as well as to improve the teaching quality of entrepreneurial foundation curriculum [22]. The findings connected to educational reform for enhancing graduate employability and human development will benefit vocational colleges, firms, representatives, instructors, and economics and trade graduates based on these findings. This research is significant for its theoretical development of finance and trade college graduate mobility, as well as an empirical assessment of education, practice, learning motivation, and personal history influences on their human development [23]. Teaching English at a vocational college is seen to be different from teaching English at a college since it is classified as English for specific purposes (ESP) and requires distinct content, methods, and tactics. The findings indicate that, for several reasons, teachers’ opinions are not always reflected in their classroom actions. Class density, time limits, incompatibility of given textbooks, massive workload, and students’ demand for consistency are all variables that contribute to the mismatch between belief and practice [24]. In comparison to the usual teaching auxiliary system, this article uses an AI module combined with information suggestions to construct an online English teaching system. The English teaching technique may be used to look at possible internal links between evaluation outcomes and other factors. This article proposes a deep learning-assisted online English teaching system, which makes use of a contemporary implementation platform to aid students in improving their English language teaching efficiency through their knowledge and personality mastery. The decision tree algorithm and neural networks were used to build an English teaching assessment implementation model based on artificial intelligence techniques [25].

3. The Linear Regression Training Model and Recommendation Framework

People used to buy things that were recommended to them by companions they respected before recently. When there was any question about the product, this was the primary method of acquisition. However, as the digital era has progressed, that circle has widened including websites that use some form of recommendation engine. This section explains how to use multilinear regression to develop a multirecommender system. By combining weights for each criterion, multilinear regression is utilized to combine the comparisons and get the complete evaluations.

3.1. The Linear Regression Training Model. The linear regression model adopts a linear relationship between the dependent variable \( y_i \) and the \( p \) vector of the regressor \( x_i \) in statistical units. The model takes the form of formula (2) due to the error variable \( \varepsilon_i \) where \( T \) represents transposition and \( x_i^T \beta \). The inner product of \( x_i \) is termed \( \beta \).

\[
 y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip} + \varepsilon_i = X_i^T \beta + \varepsilon_i, \quad i = 1, \ldots, n. \tag{1}
\]

To find the optimal \( y_i \) and \( \varepsilon_i \), an optimization algorithm is used to achieve it. There are two commonly used methods, namely, the gradient descent algorithm and the least square method [26]. The performance of the least square method is inferior to the gradient descent algorithm in optimization problems, so the gradient descent technique is selected as the optimization algorithm of linear regression in this study.

\[
\alpha_{k+1} = \alpha_k + \rho^k, \tag{2}
\]

The negative gradient direction is represented by \( s^k \), the search step length in the gradient direction is represented by \( \rho_k \), and the gradient direction is obtained by taking the derivative. The linear search algorithm is generally used to evaluate the step length.

3.2. Recommended Reading Framework. Finally, a recommendation scheme based on the linear regression training model is proposed to improve users’ reading actions and memory based on obtaining the strongest recommendation rules for vocational English reading based on the association rules recommendation algorithm and combining them with the CF recommendation algorithm based on users [27].

The overall logical structure of the English reading material recommendation scheme for English higher vocational colleges based on the linear regression training model is shown in Figure 1.

4. Analysis of English Reading Material Users in Vocational Colleges

With the need to recruit graduates who are not just paper certified but also skilled, the need for vocational education appears to be on the rise. With the rise in the unemployment problem in China, vocational education, which had been largely ignored for several years, has come into focus. The National Vocational Qualification Framework, or NVQ, as it is more often known, has opened the way for China’s students to obtain international recognition for their credentials, skills, and knowledge.

4.1. Reading Material Classification Processing. This paper expounds on the text classification of reading materials in vocational English, including the preprocessing of picture books and Fast Text classification [28]. The preprocessing process mainly includes spelling check and correction, word stem extraction, and part of speech restoration, stop word processing, and feature processing, and then the word sequence of the illustrated text is obtained. The word order at this time is represented as a one-hot representation.

\[
P(L_c | w_{01}, \ldots, w_{2m}) = \frac{\exp \left( \frac{(U^T_c(v_{01}, \ldots, v_{2m}))/2m}{\sum_{v \in V} [(U^T_c(v_{01}, \ldots, v_{2m}))/2m]} \right)}{\sum_{v \in V} [(U^T_c(v_{01}, \ldots, v_{2m}))/2m]} \tag{3}
\]
In the above formula, \( w_1, \ldots, w_{2m} \) represents the sequence of input text words, \( L_c \) represents the labels used to be predicted, and \( v_1, \ldots, v_{2m} \) represents the word vector of input words. \( \frac{v_1, \ldots, v_{2m}}{2m} \) represents the average of the input word vector, \( U_c \) represents the vector with the prediction label, and \( v \) represents the total number of terms in the text.

When constructing the Huffman tree, each leaf node is regarded as a category label, and each nonleaf node is dichotomized once [29]. The positive category probability is shown in

\[
\sigma(X_i^\theta) = \frac{1}{1 + e^{-X_i}}
\]

where 0 is the parameter of the middle node of the Huffman tree. All category tags have a path through the Huffman tree, to calculate the probability of the label \( Y_i \) corresponding to the category that the sample \( X_i \) belongs to. Formula (5) is shown as follows:

\[
P(Y_i|X_i) = \prod_{j=2}^{l} P(d_j|X_i, \theta_{j-1}).
\]  

At this point, the positive and negative category probability calculation formula is substituted into formula (5), and formula (6) can be obtained:

\[
P(d_j|X_i, \theta_{j-1}) = \begin{cases} 
\sigma(X_i^\theta), & \text{if } d_{\bar{j}} = 1, \\
1 - \sigma(X_i^\theta), & \text{if } d_{\bar{j}} = 0. 
\end{cases}
\]

The probability value obtained at this time is the probability of the category label of all kinds of vocational English reading materials.

4.2 User Behavior Collection. In the recommendation system of vocational English reading materials, users’ behavior toward vocational English reading materials reflects their different preferences for different kinds of vocational English reading materials. Data representing user behavior will be stored in the form of logs in the system background database or system files [30]. To check users’ behaviors in the system, relevant logs should be retrieved to find users’ interest in different vocational English reading materials according to their different behaviors, to make an accurate recommendation. User behavior weight is shown in Table 1:

<table>
<thead>
<tr>
<th>User behavior</th>
<th>Behavior weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Credits exchange</td>
<td>1</td>
</tr>
<tr>
<td>Read</td>
<td>0.8</td>
</tr>
<tr>
<td>Download</td>
<td>0.7</td>
</tr>
<tr>
<td>Collection</td>
<td>0.6</td>
</tr>
<tr>
<td>Share</td>
<td>0.5</td>
</tr>
<tr>
<td>Voice follow-up</td>
<td>0.4</td>
</tr>
<tr>
<td>Speed reading</td>
<td>0.3</td>
</tr>
</tbody>
</table>

4.3 Retention Rate. The calculation of the retention rate of users’ interest in vocational English reading materials is based on the combination of the Ebbinghaus forgetting curve and users’ interest in vocational English reading materials [31]. At any time in the curve, there is a holding quantity of interest matching it, denoted as \( j \), and the holding quantity function is introduced, as shown in

\[
j(t) = \frac{20^\phi}{(t_0 + c)}.
\]
The independent variable is $t$, representing the time difference between the user’s last reading and the current reading, $e$ is the base of natural numbers, $t_0 = 0.00249$. Considering the changes in users’ interest in vocational English reading materials in actual reading, the time axis magnification factor $\lambda$ is introduced.

4.4. Interest Ratio. Based on the user’s memory retention rate of vocational English reading materials and the different degrees of interest in vocational English reading materials expressed by user behavior weight, the proportion of user rate of vocational English reading materials and the diagonal factor $\lambda$ reading materials in actual reading, the time axis magnification factor $\lambda$ is introduced.

$$W_A = \frac{\left(20^\varphi/(\lambda t_{ni})^\varphi\right) \left(\sum n\varphi_{ni}\right)}{\sum \left(20^\varphi/(\lambda t_{ni} + t_0)^\varphi\right) \left(\sum n\varphi_{ni}\right)}.$$  \hspace{1cm} (8)

$W_A$ represents A English reading material in all reading material interests, the proportion of $20^\varphi/(\lambda t_{ni})^\varphi \left(\sum n\varphi_{ni}\right)$, $n$ times when reading $A$ English reading material retention, $\sum n\varphi_{ni}$ represents expressed in user behavior weight $A$ English reading material accumulation and $n$ interest degree, $\sum n\varphi_{ni}$ represents the sum of n times of interest degree of all I English reading objects represented by user behavior weight [32].

5. High Vocational English Reading Material Accurate Recommendation

The importance of English teaching materials could be understated, which is why an instructor must be able to identify the appropriate book for the teaching and learning process.

The evaluation of the material is one of the methods. This study is aimed at evaluating the language content, language skills, and themes of English teaching materials used at a vocational college.

5.1. Individual User Model. It is necessary to build a simple model for users, that is, to set up a rating table for vocational English reading materials. In the table, the number of users is denoted as $M$, and the number of reading materials in vocational English is denoted as $n$. In general, there are two ways of scoring: Boolean type and weight representation type. Boolean type is simply understood as 1 represents interest and 0 represents no interest. However, this method cannot accurately express the user’s interest degree in vocational English reading materials. Weight representation rating refers to assigning different weights based on users’ different use behaviors of vocational English reading materials [33]. User-vocational English reading score is shown in Table 2.

5.2. Interest Similarity Matching. Generally, there are three formulas for calculating user similarity, namely, the Spearman correlation coefficient, cosine similarity formula, and Pearson’s correlation coefficient. The cosine similarity formula is widely used, and its main principle is to predict the similarity between users by calculating the included angle between vectors, so it is similarly used here to calculate the parallel between two users, as shown in formula (9) for users $A$ and $B$:

$$\text{sim}(U_a, U_b) = \frac{\sum_{k=1}^{n} R_{a,k} \times R_{b,k}}{\sqrt{\sum_{k=1}^{n} R_{a,k}^2 \sum_{k=1}^{n} R_{b,k}^2}},$$ \hspace{1cm} (9)

where $R_a$ and $K$ represents evaluation vectors of all n vocational English reading materials recognized by users.

5.3. Recommendation Set Generation. According to the calculation result of the parallel between the users’ sim $(U_a, U_b)$, $K$ nearest neighbors to target user $U$ are selected. $K$ is set as 30 in this system; that is, the first 30 users are selected in descending order of cosine similarity value. The preference degree calculation formula is

$$P_{ix} = \frac{\sum_{i \in NU} \text{sim}(U_i, U_j) \times P_{jx}}{\sum_{i \in NU} \text{sim}(U_i, U_j)}.$$ \hspace{1cm} (10)

where sim$(U_i, U_j)$ represents the parallel between user $i$ and user $j$, and $P_{ix}$ represents user $j$’s score on the $x$-th vocational English reading.

5.4. Initial Recommendation List. As mentioned above, by and based on a collective filtering recommendation algorithm based on association rules after the recommendation results, we need to use a linear regression model to calculate the optimal recommended coefficient, according to the category of higher vocational college English reading interest degree of calculate the final recommendation value, according to the descending order, and according to the higher vocational English Top-n standards to produce the final recommended reading list [34]. The calculation of the optimal recommendation coefficient relies on the linear regression model. Now, it is assumed that user $U$ reading interest score of vocational English reading material $i$ under the collaborative filtering algorithm, namely, the weight is set as $R_{ui}$. The user’s final score $P_{ui}$ can be obtained from

$$P_{ui} = \alpha_{ui} R_{ui} + \beta_{ui} R_{ui}^r + e.$$ \hspace{1cm} (11)

In addition, the sum of the scoring ratio of collaborative filtering and association rules should be 1, so $\alpha_{ui}$ and $\beta_{ui}$ in Formula (11) should satisfy

$$\alpha_{ui} + \beta_{ui} = 1.$$ \hspace{1cm} (12)

5.5. Recommendation List Optimization. Through the above calculation, the user’s initial recommended list of VOCA-TIONAL English reading materials is obtained, but the list can be recommended to the user only after redundant deletion and filtering; that is, repetitive detection and deletion of vocational English reading materials are carried out to prevent the same contents of vocational English reading
materials in the recommended list. It is also necessary to filter the vocational English reading materials in the initial list [35].

To sum up, the whole process of recommendation generation of vocational English reading materials can be summarized as follows:

(1) Respectively, calculate the recommendation results under different recommendation algorithms

(2) The linear regression model is used to calculate the optimal recommendation coefficient for the results in step (1)

(3) According to the proportion of interest in various vocational English reading materials, the final recommended values of vocational English reading materials are calculated to generate the initial recommended list of vocational English reading materials

(4) Delete and filter the redundancy of initial recommended list of vocational English reading materials

(5) Generate the final recommended list of vocational English reading materials according to the top-N criterion

6. Evaluation of Simulations

6.1. Evaluation Indicators. Performance evaluation indexes in the general recommendation system mainly include three kinds, namely, accuracy rate, recall rate, and F-measure. These three indexes will be introduced in detail below:

(1) Accuracy rate: in the recommendation system, accuracy is generally used to measure the recommendation users’ interest in the content in the recommendation list [36]. Set the number of vocational English reading materials that the user likes to read as \( T \) and the length of the recommended list of vocational English reading materials as \( L \)

\[
P_u(L) = \frac{tp}{L} = \frac{tp}{tp + fn}, \tag{13}
\]

where \( tp \) represents the number of vocational English reading materials that exist in the user’s test set and his recommended list

(2) Recall rate: the recall rate is generally used to calculate the probability that the recommended vocational English reading material is in line with the user’s reading preference

\[
R_u(L) = \frac{tp}{T} = \frac{tp}{tp + fp}, \tag{14}
\]

where \( R_u(L) \) can represent the recall rate of the system

(3) F-measure: it is a comprehensive measure, a weighted average of accuracy and recall rates

\[
F_1(L) = \frac{2P(L)R(L)}{P(L) + R(L)} \tag{15}
\]

6.2. Sample Selection. In this section, practical data will be used to conduct an experimental evaluation of the English reading recommendation system in vocational colleges. During the whole experiment, the user’s reading process of vocational English with the system is simulated. To obtain the real data of future English students, 5,078 vocational English reading materials of 600 users and their behaviors were selected in this experiment, 70% of which were selected as a training dataset and 30% used for the test data set.

6.3. Recommendation Generation. Higher vocational English reading content recommendation is generated through a recommendation algorithm based on association rules which will be used for higher vocational college English reading content recommending the strongest association rules, based on combining with the CF recommendation algorithm based on user recommending the results calculated based on its own optimal recommended ratio and higher vocational English reading class’s interest degree of match; select some vocational English reading materials closest to the user’s reading preference to recommend. To verify the recommendation results of this recommendation scheme, the traditional top-N method and the recommendation method designed in this study are selected to evaluate the recommendation accuracy, and the experimental results are shown in Figure 2.

As can be seen from the above figure, the recommendation generation method of vocational English reading materials proposed in this paper is more expressive than the traditional top-N method in terms of recommendation accuracy.

6.4. Evaluation Result. Experiment and evaluate the comprehensive performance of the whole vocational English reading recommendation system. Since the recommendation scheme based on the linear regression training model adopted by the system is improved based on the traditional association rule-based and user-based CF recommendation algorithm. Therefore, in this experiment, the traditional
recommendation method is selected to analyze and compare with the recommendation method designed in this paper. The traditional recommendation technique is mainly the traditional association rule-based and user-based CF recommendation algorithm. The \( F \)-measure comparison results are shown in Figure 3.

To some extent, the recommendation method future in this study makes use of its advantages and avoids its problems, accurately expresses users’ reading interests, improves the accuracy of the recommendation generation function of picture books, and improves the overall recommendation performance of the system.

**7. Conclusion**

It is believed that utilizing reading material to improve learners’ English proficiency is highly effective. If the instructor utilizes the effective education technique, it can directly increase their reading ability. It can enable them to improve their vocabulary and enhance their command of the English language. The benefits of producing reading material for vocational high school second-grade students include enhanced reading and communication skills and improved vocabulary and language knowledge. This study proposes a recommendation scheme based on the linear regression training model. Based on obtaining the strongest association rules of vocational English reading recommendations by using the recommendation algorithm based on association rules and combining it with the user-based CF recommendation algorithm, it classifies vocational English reading texts in a more detailed and comprehensive way. Combined with user behavior data, this paper calculates the proportion of user interest in multiple aspects and vocational English reading materials, generates corresponding recommendations according to user interest, and carries out relevant experimental verification, which proves that the improved recommendation scheme has better performance in performance. In higher vocational English reading content recommendation for improvement and optimization, from the perspective of software engineering, in turn, demand analysis, general design, detailed design steps, and in the final recommendation for experiment and verification, based on the analysis of the experimental results, the conclusion, namely, the recommended scheme is put forward in this paper on the performance of the parties to perform better.

**Data Availability**

The datasets used during the current study are available from the corresponding author on reasonable request.

**Conflicts of Interest**

The author declares that he has no conflict of interest.

**References**


