Research Article

An Adaptive Learning System for English Vocabulary Using Machine Learning

Yanqiu Xu

School of Foreign Languages, Shandong Women’s University, Jinan, Shandong 250000, China

Correspondence should be addressed to Yanqiu Xu; 27047@sdwu.edu.cn

Received 18 April 2022; Revised 9 May 2022; Accepted 21 May 2022; Published 9 June 2022

Abstract

The vocabulary of a language is the collection of words used in that language. The vocabulary learning of English language plays an important role in learning English language. The expansion of learners vocabulary is linked to both their own efforts and the instruments used to guide learning. Better English vocabulary learning software can increase learner’s passion for studying as well as their learning efficiency. Therefore, this paper aims to explore an adaptive learning system for English vocabulary based on machine learning. At the beginning, it explains the primary modules of the English vocabulary adaptive learning system, as well as the system’s general design and database analysis. Afterwards, the paper looks up the AdaBoost algorithm, where the key parameter is conditional probability, which is primarily used to measure learners’ cognitive adaptation to English learning content. The fitness will be updated once learners finish the selection of English vocabulary learning content, which is known as training. The fitness will be changed gradually through training to push the relevant English vocabulary learning material and complete the adaptive learning system for English vocabulary based on machine learning. The experiments result show that the proposed learning system outperforms previous approaches in terms of learning efficiency, scientifically, and reliability.

1. Introduction

The language is a way of organized communication. The grammar of a language will be its structure, whereas the vocabulary will be its free elements [1]. Human languages are the fundamental method of communication, and they can be expressed by voice, sign, or writing [2]. Many languages, including the most frequently spoken, include writing systems that allow sounds or signs to be recorded and reactivated later [3]. Human language is different from daily communication networks because it is not reliant on a single channel of communication (sight, sound, and so on), which is extremely varied between cultures and time, as well as allows for a considerably broader spectrum of expression than other systems [4].

An important foreign language is English. The impact of English learning on academic and future job growth is significant [5]. Grammar is the backbone of language acquisition for English learners; therefore, knowing English grammar can only help them express themselves better in English; yet, if their vocabulary is limited, they will not be able to express themselves effectively even if they understand grammar [6]. The knowledge of the English language and the output of the English language are tightly linked, and vocabulary plays a significant part in the English learning level [7]. The advancement of network education and computer science has resulted in the creation of a computer English vocabulary adaptive learning system [8]. The development of a network learning system can help with the creation of a computer English vocabulary adaptive learning system [9]. The network learning system uses network resources as the primary substance of English vocabulary acquisition, the computer as an auxiliary instrument for English learning, and offers English learners with learning tools and tasks [10]. Traditional e-learning, on the other hand, focuses on the learning system, and learners must adapt to the learning system in order to gain information, which is incompatible with the educational principle of placing learners at the center; hence, learning efficiency is low [11]. To develop an adaptive learning system based on an e-learning system, the
system’s adaptability must emphasize that the learning system recognizes the differences of different learners, provides different learning resources, places the learners at the center, and adapts to the learners rather than the learners adapting to the learning system, as well as teaches students according to their aptitude [12] so that learners have their own knowledge foundation and learning aids such as cognitive style and cognitive level may effectively overcome the drawbacks of traditional computer learning and boost learning efficiency [13]. As a result, the focus of this research is on a machine learning-based adaptive learning system for English vocabulary.

The innovations of this paper are as follows:

1. Describe the major modules of the English vocabulary adaptive learning system, the overall architecture, and database analysis of the system, adopt the machine learning method, refer to the AdaBoost algorithm, and gradually adjust the fitness through training.

2. The simulation results show that compared with other learning systems, the learning system proposed in this paper has better learning efficiency and has better reliability.

The rest of the paper is organized as follows. Section 2 contains the related work to the proposed model. Section 3 is the design of English vocabulary adaptive learning system while Section 4 is the Adaptive learning system for English vocabulary based on machine learning and Section 5 is the analysis of experimental results. Finally, the paper is concluded in Section 6.

2. Related Work

With the growing importance of English vocabulary acquisition, there are varieties of vocabulary learning systems in the industry, but most of them are unscientifc, resulting in poor outcomes, and a rising number of people are studying it. Jing et al.’s analysis shows that English learning materials are abundant now, making it more difficult for students to identify appropriate resources. According to the characteristics of different learners and their learning paths, the adaptive learning system can promote to meet the learning resource difficulty and testing problems of learners. As a result, the adaptive learning system’s data organization structure is crucial, and a relational database cannot adequately reflect the significance of English knowledge. Therefore, their paper proposes the way of using ontology technology to associate with English textbook knowledge and provide learning data to the system so that learners may customize their English vocabulary learning, although this strategy has a low learning impact [14]. Deng created an English vocabulary learning system based on mobile games in order to successfully boost English learning motivation and interest. The learning system must choose English words in English textbooks based on paragraphs, difficulty ratios, and learning files. The learning activities involved in the learning system transform English vocabulary learning from cumbersome memory into game-based learning and then different types of multimedia are used to help students memorize English terminology and get more familiar with course vocabulary. By dividing students into three groups of different English vocabulary review methods for experimental comparison, student’s English level, the course contents, and teachers of all groups are roughly the same. The system assesses whether learners are still familiar with English language after playing games because the procedure is more complicated and the technique is less scientific [15]. Xiao hy, aiming at the difficulty and complexity of English vocabulary learning, combined with intelligent algorithm, proposes an English vocabulary-assisted adaptive learning system. The retrieval algorithm for English vocabulary is created, the similarity matching and image recognition technique based on word context vector are offered, and the login module and retrieval process of the learning system are supplied in order to properly implement the system. Finally, the appropriate technique of the English vocabulary-aided adaptive learning system is offered, which is integrated with the system’s interface. Although this strategy can help you acquire English vocabulary, it has a limited impact [3]. To effectively solve the problem of English learners’ vocabulary learning confusion and improve their learning efficiency and vocabulary application ability, sun FZ first examines the current state of information technology promoting English vocabulary learning and identifies relevant problems in learning resource recommendation technology, then creates English learning resources, learning characteristics, and priorities, and establishes ontology. Establish a recommendation system of English vocabulary learning resources based on knowledge map. Questionnaire survey and interviews are used to evaluate the experience of English vocabulary learning in the learning system. The relevant results show that learners are relatively satisfied with promoting the learning system. The English vocabulary learning system based on knowledge map can improve the efficiency of English vocabulary learning. However, this procedure has a scientific difficulty [4].

3. Design of English Vocabulary Adaptive Learning System

This section has been explained in different subsections such as Main Functional Modules, Overall Architecture of Adaptive Learning System, and Database Analysis of Adaptive Learning System, which are as follows.

3.1. Main Functional Modules. The adaptive learning system of English vocabulary mainly takes the research results of cognitive psychology as the construction guiding ideology and aims to realize the personalized recommendation of English vocabulary resources for learners, as well as assisting students in increasing their learning efficiency and cultivating a passion for learning [16]. The adaptive learning system is mainly divided into pre-test module, vocabulary learning module, consolidation application module, test module, and learning feedback module. Figure 1 shows the
3.1.1. Pre-Test Module. The pre-test module is primarily used to collect learners’ original data and to initialize their use of the English vocabulary adaptive learning system. The adaptive learning system’s major functions and features are to analyze learners’ characteristics and determine their original cognitive level; however, there is no useful information in the database for new users [18]. As a result, the adaptive learning system must complete the test and assessment of the initial exercise when learners use it for the first time, take the information of learners’ average question time, answer accuracy, score as the initial data, and store it in the database of learning behavior in order to accurately estimate the basic learning ability level of learners and prepare for the subsequent English vocabulary recommendations [19].

3.1.2. Vocabulary Learning Module. The key component of an adaptive learning system is the English vocabulary learning module, which primarily corresponds to the stage of information acquisition and comprehension. This module is primarily focused on the implementation of a machine learning method for assessing learner attributes. The adaptive learning system collects basic information from the behavior database, such as the average learning time of learners, the average growth of daily vocabulary, the total number of vocabulary masters, and the average time to do questions. Following machine learning, learners’ test scores at their present cognitive level are acquired and English words of significant difficulty are matched from the English vocabulary database and offered as learning materials based on the results. This function module provides a variety of English vocabulary memory methods, to achieve different learning styles of learners, because learners’ daily vocabulary will increase and their cognitive level of English vocabulary will change. After each login, learners’ learning ability will be reassessed and then matched with English vocabulary of corresponding difficulty if the necessary English vocabulary for that day has been learned; it will be marked and no longer recommended. The vocabulary that has not yet been acquired will be incorporated into the new thesaurus, and learners will be given repeated recommendations until they have learned the English vocabulary [20].

3.1.3. Consolidate Application Module. English vocabulary consolidation application module corresponds to the knowledge learning consolidation and application stage to help learners review as well as strengthen the effect of learning. This module also has the function of daily test, which can count the time for learners to do questions and give learners learning feedback, and the vocabulary recom- mender will automatically insert the incorrect English vocabulary.

3.1.4. Test Module. The test module is mainly set up to collect the characteristic information of learners and recommend English vocabulary by machine learning. If the accuracy of each test is more than 95%, it means that learners have learned all the vocabulary in the system and achieve the value of the adaptive learning system. The test content is relatively simple, and the goal is clear. It mainly focuses on the English vocabulary itself and reviews the learners’ memory of English vocabulary.

3.1.5. Feedback Module. The learning feedback module in the system can feedback learning information to learners. The learners are given feedback after the learning results; it is
beneficial for learners to reflect and grow. Figure 2 shows the growth of vocabulary per week, and Figure 3 shows the percentage of vocabulary mastered. In Figures 2 and 3, the learners can obtain feedback data from the personal center. With the help of data visualization, the percentage of mastering English vocabulary in all vocabulary, the change trend of test scores, and daily learning time are displayed by line chart and pie chart.

3.2. Overall Architecture of Adaptive Learning System. The English vocabulary adaptive learning system has simple, safe, and reliable development principles. It isolates the server from the web front end, which helps to increase learning efficiency and maintain the adaptive learning system’s fluency. The service end mainly exposes the interface and transmits it in the form of JavaScript Object Notation (JSON) data for the front end to obtain and display the interface in the call. The overall system architecture and hierarchy diagram are shown in Figures 4 and 5.

Figures 4 and 5 show that in terms of overall architecture for the English vocabulary adaptive learning system, the front end is built with sublime Text3 while the back end is created with eclipse. Tomcat is not used for web container selection in order to meet the goal of convenience and speed; thus, lightweight jetty is used instead. The application layer provides standard HTTP service interface, supports session state, and realizes web terminal docking.

3.3. Database Analysis of Adaptive Learning System. Database is the core for the English vocabulary adaptive learning system. The scientificity and simplicity of data table design as well as field design in the database will affect the speed of English vocabulary reading and writing as well as efficiency of the overall adaptive learning system. The database contains data tables of English vocabulary and learners.

The English vocabulary data sheet contains fields such as number, grade, number within grade, and difficulty value, which are represented in Table 1.

Each word in the English vocabulary has its own number. All words in the data sheet have their numbers. The numbering order is from low to high. In order to facilitate learners to learn English vocabulary, the system divides English vocabulary according to grade. The difficulty rating of words in the same grade is practically the same; thus, the grade classification is mostly dependent on the amount of English words. The grade number mostly relates to the word order in the grade. The English vocabulary difficulty value is a database property for English vocabulary that is the subject of an adaptive vocabulary learning system.

The learner data table is mainly used to record the personal information of learners, so that learners can learn based on the last login when logging in to the adaptive learning system. Six fields are represented in Table 2.

To identify the identification field of learners, the user name and password are employed. The user name cannot be registered repeatedly, but the password can be modified. The initial ability value usually refers to the ability value received by the learning system following successful registration of learners from a test of their English vocabulary ability. The level of English vocabulary gained by learners who last checked into the adaptive learning system is referred to as English vocabulary level. The field value will be changed as learners finish a level to aid English vocabulary acquisition. The test score is a field that is updated every other month. Consolidation path is the focus of adaptive English vocabulary learning in the adaptive learning system.

4. Adaptive Learning System for English Vocabulary Based on Machine Learning

The English vocabulary adaptive learning mechanism based on machine learning is composed of the AdaBoost algorithm and English vocabulary learning based on machine learning which are discussed below.

4.1. AdaBoost Algorithm. Adaptability refers to the ability of learners with various cognitive qualities to adjust to the material of English vocabulary acquisition, which is represented by conditional probability $P(B_j)$. On the premise of event $B_j$, event $A_i$ represents the probability of possible occurrence, $B_j$ represents the characteristics of class $i$.
The development tools
- WEB UI
- Template tool
- Service development tools
- Service deployment

The application layer PC
- Before using the test
- The word memory
- Review to consolidate
- A weekly test
- Information feedback

Business services

The database
- Persistence
- Cache service
- File service

WEB UI
Template tool
Service
development tools
Service
deployment
/T_he application layer PC
Before using the test
/T_he word memory
Review to consolidate
A weekly test
Information feedback

MINI UI
Vue.js

Access to the service servlet
Module The business component
MyBatis The core SQL engine

The database DB

Figure 4: Overall architecture of the adaptive learning system.

Figure 5: System hierarchy diagram.

<p>| Table 1: Glossary database. |</p>
<table>
<thead>
<tr>
<th>Field</th>
<th>Type</th>
<th>The default value</th>
<th>Primary key or not</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serial number</td>
<td>Int</td>
<td>There is no</td>
<td>Is</td>
</tr>
<tr>
<td>Level</td>
<td>Int</td>
<td>NU11</td>
<td>No</td>
</tr>
<tr>
<td>Class number</td>
<td>Int</td>
<td>NU11</td>
<td>No</td>
</tr>
<tr>
<td>Difficulty value</td>
<td>Float</td>
<td>NU11</td>
<td>No</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>Varchar</td>
<td>NU11</td>
<td>No</td>
</tr>
<tr>
<td>Paraphrase</td>
<td>Varchar</td>
<td>NU11</td>
<td>No</td>
</tr>
<tr>
<td>The phonetic symbol</td>
<td>Varchar</td>
<td>NU11</td>
<td>No</td>
</tr>
<tr>
<td>Run-on sentences</td>
<td>Varchar</td>
<td>NU11</td>
<td>No</td>
</tr>
</tbody>
</table>

<p>| Table 2: Learner database. |</p>
<table>
<thead>
<tr>
<th>Field</th>
<th>Type</th>
<th>The default value</th>
<th>Primary key or not</th>
</tr>
</thead>
<tbody>
<tr>
<td>The user name</td>
<td>Varchar</td>
<td>There is no</td>
<td>Yes</td>
</tr>
<tr>
<td>Password</td>
<td>Varchar</td>
<td>There is no</td>
<td>No</td>
</tr>
<tr>
<td>Initial capability value</td>
<td>Float</td>
<td>Nu11</td>
<td>No</td>
</tr>
<tr>
<td>Vocabulary level</td>
<td>Int</td>
<td>Nu11</td>
<td>No</td>
</tr>
<tr>
<td>Test scores</td>
<td>Float</td>
<td>Nu11</td>
<td>No</td>
</tr>
<tr>
<td>Consolidate the path</td>
<td>Int</td>
<td>Nu11</td>
<td>No</td>
</tr>
</tbody>
</table>

Table 1: Glossary database.

Mobile Information Systems 5
cognition of English vocabulary learners, and $A_j$ represents
the learners’ selection of $j$ English learning methods.
The AdaBoost algorithm is used to combine multiple weak
learners into strong learners by adjusting the probability. From the perspective of the adaptive learning system, the
algorithm first inputs the data set of English vocabulary
training into $T = \{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}$, where
$x_i \in R^n$, $y_i \in y = \{-1, +1\}$ and outputs the final strong
learner $G(x)$.

4.1.1. Initialize the Weak Learner Probability $P_1$.

$$P_1 = \left( p_{11}, \ldots, p_{i1}, \ldots, p_{iN} \right), \quad p_{ii} = \frac{1}{N}, \quad i = 1, 2 \ldots, N.$$ (1)
4.1.2. Number of Iterations $m$, $m = 1, 2, \cdots, M$. Learning English vocabulary using data from English vocabulary training, it can be obtained that the linear combination $G_m(x)$ of weak learners is expressed as follows:

$$G_m(x): x \rightarrow [-1, +1].$$

(2)

The error rate $e_m$ of $G_m(x)$ in the data set of English vocabulary training is expressed as

$$e_m = \sum_{i=1}^{N} p_{mi} I(G_m(x_i) \neq y_i).$$

(3)

The coefficient $a_m$ of $G_m(x)$ is calculated and expressed as follows:

$$a_m = \frac{1}{2} \log \frac{1 - e_m}{e_m}.$$  

(4)

Update the weak learner probability $P_{m+1}$, which is expressed as follows:

$$P_{m+1} = (p_{m+1}, 1, \cdots, p_{m+1}, I, \cdots, p_{m+1}, N),$$

$$P_{m+1,i} = \frac{P_{mi}}{Z_m} \exp(-a_m y_i G_m(x_i)), \quad i = 1, 2, \cdots, N,$$

$$Z_m = \sum_{i=1}^{N} P_{mi} \exp(-a_m y_i G_m(x_i)).$$

In the above formula, $Z_m$ represents the normalization factor, so that $P_{m+1}$ is a probability.

The linear combination of English vocabulary learning of weak learners is established, and the final strong learner $G(x)$ is expressed as follows:

$$G(x) = \sin \left( \frac{M}{m-1} a_m G_m(x) \right).$$

(6)

4.2. English Vocabulary Learning Based on Machine Learning.

The goal of an adaptive learning system is to tackle the problem of learners’ fitness being dynamically updated when they acquire an English vocabulary. Refer to the above AdaBoost algorithm formula to simplify it, which is specifically expressed as follows.

Take the data set of learners’ selection of English vocabulary as the training data set and record it as $T = \{(x_1, y_1), (x_2, y_2), \cdots, (x_N, y_N)\}$ where $x$ is the type of English vocabulary learning content, $y$ is the result of selection, $i$ is the content of English vocabulary learning, $n$ is the number of types of English vocabulary learning content, $x = \{1, 2, \cdots, i, \cdots, n\}$, and $y = \{+1, -1\}$.

Record $P(c_i | s = s_0)$ as $P_i$. At present, the fitness is

$$P = (P_1, P_2, \cdots, P_n, \cdots, P_n).$$

If it is initialized, $P_i = 1/n$; otherwise, it is the fitness obtained from the last operation.

The adaptability of English vocabulary learning after updating is $\bar{P}$, $\bar{P} = (P_1, P_2, \cdots, P_n, \cdots, P_n)$.

The error rate $e$ of English vocabulary learning is calculated and expressed as follows:

$$e = \frac{1}{n} \sum_{i=1}^{n} x_i.$$  

(7)

The coefficient alpha is calculated and expressed as follows:

$$\alpha = \frac{1}{2} \log \frac{1 - e}{e}.$$  

(8)

Get the updated $\bar{P}_i$: $Z$ represents the normalization factor.

$$\bar{P}_i = P_i \exp(-\alpha y_i (-1)),$$

$$Z = \sum_{i=1}^{n} P_i,$$

$$\bar{P}_i = \frac{P_i}{Z}.$$  

The fitness $\bar{P}$ of English vocabulary learning after updating is obtained, and the above process completes the adaptive learning of English vocabulary.

5. Analysis of Experimental Results

In order to verify the performance of the adaptive learning system for English vocabulary based on machine learning proposed in this paper, it is proved by the following experiments. The experimental environment is shown in Table 3.

Table 4 shows the comparison between the adaptive learning system for English vocabulary based on machine learning and the English vocabulary adaptive learning system based on fuzzy logic. The average score formula of system test score is as follows:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} x_i.$$  

(10)

The data in Table 4 are obtained after phased learning. The learning performance of the two methods has improved, but compared with the learning performance of the two methods, the adaptive learning system based on machine learning proposed in this paper has significantly higher performance than the adaptive learning system based on fuzzy logic, indicating that the reliability of the method proposed in this paper is higher.

Figure 6 shows the comparison of learning efficiency between the method proposed in this paper and the method based on fuzzy logic. It can be seen from Figure 6 that with the continuous increase in the scale for learning resources, the learning efficiency of the English vocabulary adaptive learning method based on fuzzy logic has increased. However, the overall learning efficiency is relatively lower than that of the adaptive learning method for English vocabulary based on machine learning proposed in this paper, indicating that the stability of the method proposed in this paper is good, and it can effectively improve the learning efficiency of learners.
Figure 7 shows the scientific comparison between the vocabulary adaptive learning system for English based on machine learning and the English vocabulary adaptive learning system based on fuzzy logic. Through the analysis of Figure 7, it can be seen that with the gradual increase in English vocabulary, the scientific rationality of the two adaptive learning systems gradually increases, but the English vocabulary adaptive learning system based on machine learning has a higher scientific rationality than the fuzzy logic system.
In this paper, the system based on fuzzy logic fluctuates greatly. In comparison to previously, when the English vocabulary hits 600, it decreases. The method proposed in this paper has been showing a gradual upward trend with the increase in vocabulary; it demonstrates that the machine learning-based adaptive learning system’s integrity is generally consistent and may successfully boost students’ interest in learning English in order to improve their academic performance. Figure 8 shows the effect comparison between the adaptive learning system for English vocabulary based on machine learning and the English vocabulary adaptive learning system proposed in literature [3, 4].

With the gradual increase in vocabulary mastered by learners, as shown in Figure 8, the overall learning effect of the English vocabulary adaptive learning system proposed in literature [3] and literature [4] is poor. It is indicated that learners’ learning efficiency is low and the amount of vocabulary mastered is small, and therefore adaptive learning system’s significance is lost. Because the main parameter of machine learning method is conditional probability, it mainly measures the adaptability of learners’ cognitive characteristics to English learning content, and it can also help improve learning efficiency. The effect of the adaptive learning system for English vocabulary based on machine learning proposed in this paper increases with the increase in English vocabulary.

6. Conclusions

In the face of massive information resources and learning content, it is difficult for learners to find the best learning materials suitable for themselves, thus wasting time and cost. The adaptive learning system is mainly characterized by the recommendation of personalized English learning resources and the guidance of learning path, which solves the problem of cognitive overload of learners to the great extent. The purpose of the paper is to help learners find appropriate English vocabulary resources, guide them down a scientific learning path, and assist them in improving the efficiency of their English vocabulary memory, achieving learning goals, optimising English vocabulary learning methods, and improving their learning ability. It discusses the English vocabulary adaptive learning system’s main modules, as well as the system’s overall architecture and database analysis. It employs the AdaBoost algorithm, with conditional probability as the primary parameter, to assess learners’ cognitive adaptation to English learning content. When learners complete the exercise, or selection of English vocabulary study resources, their fitness will be upgraded. The fitness will be gradually improved through training in order to push appropriate English vocabulary learning material and complete the machine learning-based adaptive learning system for English vocabulary. The results of the experiments reveal that the proposed learning system exceeds earlier systems in terms of learning efficiency, scientific validity, and consistency.

Data Availability

The data pertaining to this research are included for publication of this work.

Conflicts of Interest

The author declares that there are no conflicts of interest for publication of this paper.

References


