Research Article

Construction of Enterprise English Adaptive Learning Platform Based on Big Data Analysis

Chen Luo

Faculty of Foreign Languages and Business, Jiaozuo Normal College, Henan, Jiaozuo 454000, China

Correspondence should be addressed to Chen Luo; luochen1984@jzsz.edu.cn

Received 15 March 2022; Revised 28 April 2022; Accepted 23 May 2022; Published 4 July 2022

Academic Editor: Sheng Bin

Copyright © 2022 Chen Luo. 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.

With the rise of “Internet+” as a national strategy, traditional industries have accelerated their integration with the Internet, and the education industry has also begun to seek innovative development. The fast-paced life of modern people makes time for learning more fragmented, and the high pressure of life makes people want to avoid excessive pressure from learning. These give opportunities and advantages to the development of online education, which is more adaptable to fragmentation and lightweight. Online learning effectively solves the drawbacks of imbalance between supply and demand of offline education by solving the problems of time, distance, capacity, and scarcity of educational resources. In this paper, we take H International Education Group as the research subject. The paper combines the practice of precise English teaching theory and big data technology analysis tools to study the optimization of English teaching strategies in the K12 English training product market of H International Education Group. The materials prevalent in the English vocabulary market today are “one-size-fits-all,” presenting the same vocabulary resources to all learners without taking into account the differences in learning styles, learning preferences, and original cognitive structures of different learners.

1. Introduction

The traditional culture of Confucianism advocates the traditional Chinese educational concept of “learning is superior,” which has made the Eastern people more enthusiastic about the education of only children [1]. Education and upbringing are organic and complementary and have been the most pressing concerns of Chinese parents since the beginning of time [2]. The way of educating and guiding students to absorb knowledge has also undergone a fundamental change with the rapid development of globalization and information technology [3]. However, the core educational objective of “optimizing students’ learning styles and improving the rate of knowledge absorption” remains the same [4]. Recently, the Decision of the Central Committee of the Communist Party of China on Several Major Issues of Comprehensively Deepening Reform clearly proposed to accelerate the construction of an effective linkage mechanism that actively uses modern information technology to gradually expand the information coverage of high-quality higher education resources [5]. It further clearly foretells that the traditional vocational education management mode in China will further gradually integrate with the management mode of vocational education in depth with the rapid development of technology in the information age in the future, thus gradually forming a new concept of vocational education development and a series of new innovative vocational education management modes and systems [6–9].

With the massive amount of information data, practitioners are required to explore new processing modes in the big data system in order to obtain effective information from it, use it, and analyze it with data mining and machine learning methods and other technical [10]. It can be seen that collecting massive and diversified kinds of information data and using technical means to obtain and predict the statistical laws of things will assist in improving the ability of making various decisions by the state, enterprises, and individuals [11]. For example, in the field of business English teaching, the key to profitability lies in improving the
“processing and analysis” ability of data through the “processing of data” to realize the profitability of business [12–15]. The key to profitability in business ELT, for example, is to improve the “processing and analysis of data” and to “add value to data or ELT profits” through the “processing of data.” Big Data for ELT is an important segment of big data technology system in the process of commercial ELT activities [16]. It contains not only the definition of data dimensions, access paths and tools, hardware and software infrastructure, data cleaning, storage hardware and software system, analysis algorithms, software tools, and other necessary elements in the general big data system, but also additional definition and strict requirements for commercial ELT activities [17]. The data dimension system, user behavior profile system, customer consumption preference system, product and service intelligent recommendation system, market opportunity and risk warning system, and other application frameworks refined by the English teaching management framework research can be quickly applied to the matching industry and digital intelligence English teaching activities, optimize the English teaching input-output ratio (ROI), and effectively reduce the cost of English teaching in enterprises [18–20].

The value of business application of big data is of particular interest, as it can assist companies and enterprises to segment customer groups and provide high-quality customized services [21]; it can simulate the real environment based on big data to discover new customer needs; it can reduce the basic service overhead of companies and enterprises and save human resources. This paper introduces the basic principles of data mining statistical methods: correlation analysis, factor analysis, and cluster analysis, where correlation analysis uses Pearson correlation coefficient and Spearman correlation coefficient to calculate the correlation between variables, and cluster analysis uses k-means analysis method. The application of data mining to the analysis of students’ performance is emphasized; the correlation analysis is used to study the degree of correlation between different subjects in junior and senior high schools; and the factor analysis is used to study the different learning abilities of students at the same level, to provide a basis for students and parents to select courses in the new college entrance examination “3 + 3,” to provide schools and teachers with class placement suggestions, and to diversify the evaluation of students’ learning abilities. The application of data mining to the evaluation of students’ comprehensive quality, based on the existing evaluation method of “good education” in Xi’an Huian Middle School, provides the correlation between comprehensive quality evaluation and college entrance examination results and provides data support for comprehensive quality evaluation.

First, this paper systematically explains the current situation of the education market at home and abroad and proposes a research idea of how to optimize the English teaching strategy of H International Education Group by using precision English teaching theory. Second, the third-party database of the education industry and machine learning methods are used to analyze the four dimensions of the current situation of H International Education Group’s English teaching strategy: product customer positioning, brand market positioning, English teaching strategy positioning, and organizational system culture. Again, the combination of big data platform tools and machine learning algorithms (k-means) and other technical means is fully utilized to innovatively propose a high-end channel precision English teaching model based on channel customer portrait identification in the K12 education industry.

2. Data Collection and Preprocessing for Corporate English Learning Samples

2.1. Data Preprocessing Methods and Steps. The process of establishing and implementing accurate English teaching requires the enterprise or team to have the foundation of various production elements, from the basic elements to the design and implementation of various business management systems, strategies, and tactics at the upper level, and only by meeting the necessary digital and agile infrastructure conditions can the business team clearly understand the business advantages and disadvantages faced by the enterprise, accurately grasp the dynamics of market changes, and adjust the strategies and tactics according to the state of market competition. In order to enable the business team to clearly understand the company's business advantages and disadvantages, accurately grasp the market dynamics, and adjust its strategies and tactics according to the competitive situation, digital and agile infrastructure is needed. Focusing on the case of H International Education Group's high-end channel English teaching, the group's management quickly identified the necessity and urgency of establishing accurate English teaching, relying on the CRM and database reports of the marketing and sales software already deployed and running in the group, as well as the introduction and training of talents for accurate English teaching, high-end channel English teaching, big data analysis, and online and offline operations. The urgency and systematic approach of the establishment process were quickly identified. In the group's overall corporate strategy, the implementation budget and strategic height of the precise English teaching strategy and high-end channel English teaching management methods were fully defined, which ensured that the difficulties in the preparation of various infrastructures were overcome and quickly implemented from the financial and administrative perspectives, and also ensured that the group's channel customers' product services reflected the service commitment to rapid improvement in a short period of time.

Figure 1 is a line graph of students’ vocabulary acquisition over time in the system. The horizontal coordinate is time, and the vertical coordinate is the total number of vocabulary items mastered by the learner. If the number of vocabulary items mastered is positively correlated with time and increases steadily with time, the learners are highly motivated to learn during that period of time; if the number of vocabulary items mastered is negatively correlated with time and stays the same or decreases with time, the learners are depressed and lack motivation to learn during that period of time. The slope of the line can also indicate the
degree of motivation. The tilt of the line can also indicate the learning status, and the larger the positive tilt angle, the more efficient the student in that period of time, and the opposite is true. By looking at the graphs, learners can analyze the problems that exist over time and find effective solutions.

H International Education Group attaches more strategic importance to the collection and use of business production data than the domestic K12 education market in general. The online teaching part of its K12 curriculum education system, the Internet classroom department, and the offline multimedia interactive classroom department, which are the characteristics of H International Education Group’s education products, will naturally produce various business data, but these business data and sales department data, customer management data, and market dynamic data are not naturally related. However, these business data are not naturally related to the sales department data, customer management data, and market dynamic data, which need to form a unified, highly related, and centrally accessible enterprise data pool through a series of data specifications, business specifications, and system management specifications, which we collectively call “Enterprise Big Data Platform Strategy DMP Strategy.” Enterprise Big Data Platform Strategy DMP Strategy includes the core components shown in Table 1.

The management of H International Education Group has strategically invested in and designed a corporate DMP strategy to quickly integrate and optimize the various organizational management structures and core departments such as product, operations, and marketing ELT through an enterprise data platform product tool set for digital business flow. To realize the business foundation of digital enterprise and digital English teaching, DMP (Data Management Platform) is the core prerequisite for the rapid establishment of an accurate English teaching system. Due to the history and positioning of the telecom industry in China’s domestic communication development, along with its innovative application in the national production process and various social productivity and business logic, the telecom industry has long accumulated customer business data with high-quality big data characteristics such as high real name, high geographical logic, high track tracking, high consumption habits, and customer industry channel differentiation. At the same time, due to the high mobility of telecom equipment,

---

### Table 1: Core elements of enterprise big data platform strategy.

<table>
<thead>
<tr>
<th>Platform</th>
<th>Core elements</th>
<th>Feature dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Product business data</td>
<td>Design standard system</td>
</tr>
<tr>
<td>2</td>
<td>User data tagging</td>
<td>System design</td>
</tr>
<tr>
<td>3</td>
<td>Business data metrics dimension</td>
<td>Labeling system design</td>
</tr>
<tr>
<td>4</td>
<td>ELT dimensions</td>
<td>Big data system design</td>
</tr>
<tr>
<td>5</td>
<td>ELT OA/CRM unified data</td>
<td>Indicator set design</td>
</tr>
<tr>
<td>6</td>
<td>Big data computing</td>
<td>Hardware system design</td>
</tr>
<tr>
<td>7</td>
<td>Industry third-party database</td>
<td>Data label unification design</td>
</tr>
<tr>
<td>8</td>
<td>Big data analytics</td>
<td>Algorithm standard system</td>
</tr>
<tr>
<td>9</td>
<td>Big data business</td>
<td>Reporting form and criteria</td>
</tr>
<tr>
<td>10</td>
<td>Big data business team</td>
<td>Business KPI</td>
</tr>
</tbody>
</table>
the massive and real user geographic location movement system accumulated by telecom customer database can be combined with GIS business geographic map data system to obtain a highly accurate offline business trajectory heatmap. The K12 education industry has many offline business scenarios, and the third-party database of telecommunication can provide accurate heat distribution of potential customers’ location movement trajectories. The special advantage of such a third-party English teaching database can quickly help H Education Group to make targeted adjustments to the deployment of its offline English teaching location layout, especially by using big data such as GraphQL graph databases like Neo4j. In particular, using big data deep learning mining technology such as GraphQL databases like Neo4j, we can use the desensitized big data information in the telecommunication database and the intermediate analysis results to match the customers with English teaching channels, and quickly locate the contact information and English teaching scenarios of bank financial customers in Sichuan and even western region, such as credit card English teaching scenarios, bank English teaching scenarios for youth K12 financial products, bank English teaching scenarios for education and savings products, and insurance-based financial products English teaching scenarios. The results are cross-referenced with H International Education Group’s own channel English teaching data system to be verified, so as to accurately push the channel customer list and English teaching strategy planning to the channel English teaching team of the group and the campus product and service infrastructure.

2.2. Determination of Neural Network Structure. With the huge amount of information data, practitioners are required to explore new processing modes in the big data system in order to obtain effective information from it and use the information, such as data mining and machine learning methods and other technical means to analyze it. It can be seen that, by collecting massive and diverse kinds of information data and adopting technical means to obtain and predict the statistical laws of things, it will assist in improving the ability of making various decisions by the state, enterprises, and individuals, as shown in Figure 2. For example, in the field of business English teaching, the key to profitability lies in improving the "processing and analysis capability" of data, and through the "processing and processing of data" to achieve business profitability. The key to profitability in the field of business English language teaching, for example, is to improve the "processing and analysis of data" and to "add value to the data or to the profitability of English language teaching" by "processing the data."

Big Data for English teaching is an important segment of big data technology system in the process of commercial English teaching activities. It contains not only the definition of data dimensions, access paths and tools, hardware and software infrastructure, data cleaning, storage hardware and software systems, analysis algorithms, software tools, and other necessary elements of the general big data system but also additional definition and strict requirements for commercial English teaching activities.

The code that allows the back-end to project a student’s predicted grade based on the data queried in the database is shown below.

```java
public double getScore(int testTime[], int studyTime[], int totalNum[], int addNum[]) throws Exception{
    double score = 0;
    for(int i = 0; i < testTime.length; i++){
        for(int j = 0; j < studyTime.length; j++){
            for(int k = 0; k < totalNum.length; k++){
                for(int l = 0; l < addNum.length; l++){
                    int singleScore = 0;
                    DataRow datarow = dao.findWeightByLevel(testTime[i], studyTime[j], totalNum[k], addNum[l]);
                    if(datarow.msgCode != 0){
                        switch(datarow.scoreLevel){
                            case ScoreLevel.one:
                                singleScore = 20 * datarow.weight;
                                break;
                            case ScoreLevel.two:
                                singleScore = 40 * datarow.weight;
                                break;
                            case ScoreLevel.three:
                                singleScore = 60 * datarow.weight;
                                break;
                        }
                        score += singleScore;
                    }else{
                        throw new BusinessException();
                    }
                }
            }
        }
    }
    return score;
}
```

2.3. Breakdown of Scenarios. When the user lifecycle is maintained at a certain level, the increase of user value per unit of time can lead to an increase in total revenue. This can be achieved by adding more value-added services, meeting the needs of user segments, etc. H International Education Group’s differentiated services designed to meet the needs of high-end channel customers can effectively increase the frequency and quantity of service purchases by these channel customers and gain a higher share of K12 education product sales from service segmentation, which increases the LTV of product customers.
Using the data burial points of our own Internet tools and the data collection and clustering analysis of third-party SDK tools, we can draw user portraits in advance and then match the corresponding marketing paths and marketing forms and contents for different people on the basis of user tags, so as to meet the target customer groups in layers of refinement and satisfy the paying consumers of educational products, especially those with high LTV, high customer customization attributes, and high service requirements. VVIP customer group users with high LTV, high customer customization attributes, and high service requirements provide the right marketing scenarios and means, as shown in Table 2, to easily achieve efficient sales conversion of marketing costs.

By selecting and comparing various high school English vocabulary difficulty classification systems, the author finally selected the high school English curriculum standard vocabulary list compiled by the Education and Research Center as the reference standard for the vocabulary ranking system of this study. Figure 3 is a partial screenshot of the table. The vocabulary list contains about 3500 high school English vocabulary words, which are divided into five levels: 1, 2, 3, 4, and * according to the frequency of vocabulary use. 1 to 4 levels of vocabulary words decrease in frequency step by step, the difficulty coefficient decreases step by step and difficulty increases, while those marked with * are some super vocabulary words with the smallest difficulty coefficient and the greatest difficulty. However, in terms of the increase in mean score, the experimental group was slightly higher than the control group, indicating that the use of the adaptive vocabulary learning system helped learners’ vocabulary memory to a certain extent. With in-depth observation, it was found that the maximum scores of the pretest and posttest groups changed by approximately the same amount, while the minimum scores of the experimental group increased significantly more than those of the control group.

### 3. Results and Analysis

#### 3.1. Learning Achievement Test

The results of the test can best reflect the students’ vocabulary mastery. Since the questions of the test module of this system target all the vocabulary required in the college entrance examination syllabus, and the difficulty level of the test questions remains the same, the results of the pretest and posttest of the experiment can well reflect the changes of the learners’ mastery level. As shown in Figure 4, in terms of mean scores, since the experimental subjects were divided into two groups of relatively equal learner ability based on the pretest scores, the mean scores of the pretest of the experimental group and the control group were basically the same. According to the formula for finding the average score, the pretest control group scored 52, with a maximum score of 82 and a minimum score of 24, and the experimental group scored 54, with a maximum score of 84 and a minimum score of 20. The mean score of the posttest results was 68 for the control group, with a maximum score of 92 and a minimum score of

---

**Table 2: Adaptive platform behavior weighting analysis.**

<table>
<thead>
<tr>
<th>Weight ranking</th>
<th>Behavior templates</th>
<th>Analytical tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The needs of core users of enterprise products and services</td>
<td>Behavioral digital portraits</td>
</tr>
<tr>
<td>2</td>
<td>Industry-wide marketing channels</td>
<td>This data dynamic matrix</td>
</tr>
<tr>
<td>3</td>
<td>Customer lifecycle value tracking</td>
<td>Data tools</td>
</tr>
<tr>
<td>4</td>
<td>Marketing big data</td>
<td>Analytical forecasting tools</td>
</tr>
<tr>
<td>5</td>
<td>Competitive word-of-mouth detection tool</td>
<td>Demand detection tool</td>
</tr>
<tr>
<td>6</td>
<td>Service solutions</td>
<td>Friendliness of experience</td>
</tr>
<tr>
<td>7</td>
<td>Internet content and social networks</td>
<td>Platform traffic tools</td>
</tr>
</tbody>
</table>

---

**Figure 2: Enterprise English big data analysis process.**

---

**Figure 3: Table 2 screenshot.**

---

**Figure 4: Enterprise English big data analysis process.**

---

**Scientific Programming 5**
38, and 74 for the experimental group, with a maximum score of 94 and a minimum score of 52. To address this situation, the variance of the scores of the two groups was then analyzed. The variance represents the degree of dispersion of the values, and the larger the variance, the larger the gap between the values and vice versa, the smaller the gap.

From the above data, we can conclude that after one month of study, both the experimental group and the control group had a relatively large increase in their scores, which indicates that the students were in a good learning state and put more effort into vocabulary memorization, and the students’ full commitment also ensured the reliability and accuracy of the experimental results.

According to the variance formula, the variance of the control group was 397.04 and 268.75, and the variance of the experimental group was 392.63 and 138.39. From these two sets of data, it can be seen that due to the deliberate division of the ranking before the test, the two groups of students’ scores changed by roughly the same amount, and the fluctuations of the scores of the two groups were similar, so the variance values were close. After one month of experimental testing, the variance of the experimental group was significantly smaller than the variance of the control group, which indicates that the score gap between the learners in the experimental group is smaller than that of the control group. In order to clarify the reason, the author counted the posttest scores of the experimental group and the control group according to the number of score bands, and the scale diagram shown in Figure 5 can be derived. The system satisfies the condition that each individual has different components of intelligence through personalized vocabulary resource recommendation, and reduces the pressure of intrinsic cognitive load. Through various memory aids such as pictures and pronunciation, the system creates a memory situation for learners, allows learners’ multiple intelligences to work together, cultivates creative thinking, accelerates the operation of intelligences, and improves memory ability, while also reducing the pressure of extrinsic cognitive load and lowering the difficulty of memory.

From the above graph, we can find that the number of people in the low score range (20–60) of the experimental group is much smaller than that of the control group, while the number of people in the 60–80 score range is higher than that of the control group, which indicates that the mean score of the experimental group is higher than that of the control group. From the above graph, we can find that the number of people in the low score range (20–60) of the experimental group is much smaller than that of the control group, while the number of people in the 60–80 score range is higher than that of the control group, which indicates that the mean score of the experimental group is higher than that of the control group. From the above graph, we can find that the number of people in the low score range (20–60) of the experimental group is much smaller than that of the control group, while the number of people in the 60–80 score range is higher than that of the control group, which indicates that the mean score of the experimental group is higher than that of the control group. From the above graph, we can find that the number of people in the low score range (20–60) of the experimental group is much smaller than that of the control group, while the number of people in the 60–80 score range is higher than that of the control group, which indicates that the mean score of the experimental group is higher than that of the control group.
reduces the influence of intrinsic cognitive conformity in cognitive load, so that resource information can be maximally understood and processed by learners to improve learning efficiency and achieve learning goals.

By designing and distributing a questionnaire, I collected data from the learners to analyze their emotional attitude towards the system and to draw conclusions that could be used as a basis for improving the system. The author distributed 25 questionnaires to learners who had experience in using the system, and 25 valid questionnaires were collected; the statistical information is shown in Figure 6. To sum up the above data analysis, we can conclude that the adaptive vocabulary learning system is helpful for learners, and the most suitable group of learners is those who have weak English learning ability and insufficient previous learning experience. This also supports the theoretical basis of this study by showing that learners’ learning resources are assigned in a personalized way and the difficulty of vocabulary is in line with learners’ previous cognitive experience.

From the statistical information, most of the learners agree that the recommended vocabulary resources are necessary and their learning performance has improved significantly through the recommended resources. In terms of the recommended resources, most of the learners thought that the vocabulary was not too difficult to learn and was within the acceptable range, which verified the reasonableness of the system’s vocabulary recommendation, which could fit the learners’ original cognitive structure, achieve personalized learning, save time cost, and ensure the efficiency of learning. The effectiveness of this part mainly lies in the use of fuzzy logic reasoning mechanism, which effectively analyzes the learning characteristics of learners and recommends reasonable and reliable learning resources for individuals with different intelligence components, effectively reducing the intrinsic cognitive load and allowing learners to improve their learning efficiency and cultivate their learning interest. The system’s interface presentation mainly benefits from the use of technical means of data visualization.

Referring to the visual evaluation criteria, as shown in Figure 7, according to the information collected from the questionnaire, it can be seen that in terms of comprehensibility, all learners agree that the system presents information clearly and distinctly, that they can face the learning process and get feedback information directly, and that the comprehensibility of the system is good. In terms of visual decorativeness, most of the learners affirmed the system’s auxiliary functions of vocabulary pictures, pronunciation, and related vocabulary listing, and the system is helpful for helping learners to construct a context, expand their imagination, and help them remember vocabulary. In terms of metaphorically, learners also gave consistent affirmation to the presentation of the learning process and data feedback, believing that the charts can well reflect the learning dynamics of different learning stages and help learners understand their learning status in a timely manner. In terms of aesthetics, the system was not well accepted, with 60% of the learners giving a negative opinion. The content of the questionnaire includes two aspects of investigation. On the one hand, whether the system realizes the personalized allocation of resources by using fuzzy logic. On the other hand, based on the five evaluation criteria of data
visualization, consider the rationality of using data visualization technology.

As shown in Figure 8, the system’s use of two technical tools, fuzzy logic and data visualization, has achieved certain results, which helps learners to efficiently memorize vocabulary, improve their memory ability, and accomplish their learning goals and also proves the significance of using multiple intelligence theory and cognitive load theory as the theoretical basis for guidance. At the same time, the system summarizes the related vocabulary system of specific words, which increases the associated cognitive load to a certain extent and helps learners expand their thinking, acquire learning skills, master efficient learning methods, and improve their learning ability.

For texts with chronological order, their content has typical chronological characteristics, such as a story that changes according to the sequence of time development or a tracking news story that advances according to the increment of time. The learning activity is also a dynamic behavior that steadily proceeds with the passage of time, and allowing learners to directly witness their own learning trajectory plays a crucial role in learning feedback. The data dimension system, user behavior profile system, customer consumption preference system, product and service intelligent recommendation system, market opportunity and risk warning system, and other application frameworks refined by the English teaching management framework research can be quickly applied to the matching industry and digital intelligence English teaching activities, optimize the English teaching input-output ratio (ROI), and effectively reduce the cost of English teaching in enterprises, as shown in Figure 9.

4. Conclusion

The development of technology has brought about changes in the means of education, and in the information age, online learning has become one of the most popular ways of
learning. Faced with the explosion of information resources and massive learning contents online, learners cannot easily find the optimal learning materials for themselves, and this also greatly wastes their time costs. This study is based on multiple intelligence theory and cognitive load theory, using the two technical means of fuzzy reasoning mechanism and data visualization to ensure the rationality of recommendation and the scientificity of presentation, focusing on high school English vocabulary learning. The study aims to recommend vocabulary resources that fit learners’ characteristics; guide scientific and effective learning paths; provide timely and accurate learning information; and help learners improve their vocabulary memory efficiency, fulfill their learning purposes, examine their learning attitudes, adjust their learning styles, and enhance their learning abilities. After certain experimental validation, the main result of this study is as follows: the rationality of vocabulary recommendation is ensured by means of fuzzy reasoning mechanism.

In this study, the layout of the interface and the color scheme of the system still need to be improved, and this is an area that needs to be addressed in the future. In terms of memorability, most of the learners gave a positive attitude, indicating that the system has a certain role in helping learners to improve their memory effect.

**Data Availability**

The data used to support the findings of this study are available from the author upon request.

**Conflicts of Interest**

The author declares no conflicts of interest.

**Acknowledgments**

The study was supported by the Faculty of Foreign Languages and Business, Jiaozuo Normal College.

**References**


