Application Data Mining Technology in Construction of College English Blended Teaching Mode

Wei Wei*
Wuhan International Trade University, Wuhan City, Hubei Province, 430000, China

*Corresponding author: 22675666@qq.com

Abstract. With the opening up of China, the communication with the international community is increasing, and the technical talents who can adapt to the international trade are more and more eager. But at the same time, there are still some problems in English Teaching in higher vocational colleges, such as the lack of teaching time, the uneven English foundation of students due to different sources, and the low enthusiasm and fear of difficulties in English learning. Based on the practice design of mixed teaching mode, this paper analyzes the statistical results. According to the results, SIG. Value is 0.238, greater than 0.05, and sig. (bilateral) value is 0.047, less than 0.05. According to this significance level, we can know that there is a significant difference between the two classes. Therefore, compared with the traditional English teaching, SPOC teaching mode based on blended learning theory can improve students' overall English level and learning efficiency.

Keywords: Data Mining, College English Teaching, Blended Teaching, Data Analysis

1. Introduction
In 2012, the Ministry of Education issued the "ten year development plan of education informatization (2011-2020)" , which mentioned: to promote the integration of information technology and teaching, provide students with intelligent learning environment, provide high-quality education and teaching resources, use modern information and network technology to carry out heuristic, exploratory and interactive teaching in Colleges and universities, promote formative evaluation, and strive for the development of higher education We should establish a new student-centered teaching mode, promote the cooperation between schools through the network, and comprehensively improve the level of information-based teaching; vigorously promote students to use network information technology to carry out autonomous learning and cooperative learning; let students form the habit of using network information technology to learn, cultivate their personality and expertise, and improve the learning effect; enhance students' ability in network learning The ability to find and analyze problems in the information environment, and finally solve problems. This also points out the development direction for our current education reform [1].

With the development of globalization, the pace of China's "reform and opening up" is faster and faster, and the relationship between China and the world is becoming closer and closer. China's policy
of one belt, one road, brought about 6 trillion US dollars in trade volume from 2013 to the end of 2018. The policy of "opening to the outside world" urges more and more Chinese enterprises to go abroad, and also forces enterprises to put forward higher requirements for industrial upgrading and personnel quality, especially for the demand for compound talents who are "proficient in both English and technology". Under the above circumstances, higher vocational English teaching is also facing higher challenges. But at the same time, due to the influence of traditional ideas, the source of higher vocational students is not very good, the hierarchical structure of higher vocational education system is not perfect, the lack of teachers and other reasons, higher vocational education is also facing a variety of problems, and these problems also exist in Higher Vocational Public English teaching. Public English is a required course for non English Majors in higher vocational colleges. According to the regulations of the Ministry of education, vocational college students are required to master certain English knowledge and skills, have the basic abilities of listening, speaking, reading, writing and translating, and be able to understand English professional materials with the help of translation tools. In practice, the author makes a statistics of 16 Vocational Colleges in Hangzhou, and finds that the longest colleges offer three semesters of Public English, a total of 124 class hours; while most colleges, including the schools where the author works, have only one academic year, including two semesters, about 80-100 class hours; even some colleges have only one semester of English, not enough To 50 class hours. This is far from the class hours stipulated by the state. The reduction of teaching time directly leads to the shortage of students' English learning time. So what we need to solve is how to make efficient use of extracurricular learning time in addition to the limited classroom time [2].

Based on the summary and analysis of relevant theoretical literature at home and abroad, aiming at the existing problems in English Teaching in vocational colleges, this study takes blended learning as the theoretical basis, combined with SPOC teaching form, and verifies the role of this blended teaching mode in practical teaching through concept elaboration and practical design and analysis.

2. Related Concepts

2.1 Data Mining
Data mining is simply to extract valuable and hidden information from a large number of data through analysis technology, establish corresponding model, and use this model to help people make correct decisions. Extracting valuable information is the key of data mining, but the valuable information is relative, and different fields need different information. For example, in the banking industry, they need to extract customer characteristics from numerous customer information, classify each customer, and excavate customers' credit ability or consumption ability. This idea is used in the bank's precision marketing. By analyzing which kind of customers are easy to buy financial products, they can develop such customers well and form a stable source of profit. In addition, in the process of lending, by analyzing the characteristics of previous default and non default customers, we can establish a credit scoring system for customers and guide credit officers to lend. Data mining is not the ability of simple knowledge mining, it can visualize the analysis, through some charts or shapes, such as decision tree technology, by drawing a very easy to understand tree diagram, the bifurcation of each branch represents a situation, which is the most intuitive analysis method for analysts, and the most convenient for decision makers You can watch the analysis results, and stimulate the inspiration of decision-makers through visual impact to make the best decisions [3].

2.2 Basic Process of Data Mining
Now with the gradual development of big data technology, the application process of data mining technology is gradually mature, applied in many fields, and the analysis process is gradually standardized. After studying a large number of data mining cases, it is found that the basic process of data mining can be divided into the following steps [4].

Data collection: get the original data needed by data mining. There are many methods of data collection, the most commonly used is to get and crawler from the database. For example, the banking
industry will store customers' data, which is a very good data source; when it comes to analyzing the fundamentals of a company, it can capture the enterprise information it needs through crawler technology [5].

Data preprocessing: if you directly use the original data for modeling, there may be some problems. For example, some variables in the data contain a large number of missing values, which will seriously affect the final result; some variables contain a large number of duplicate values, which has little significance for modeling; some data are not numerical, and need to be characterized.

Data cleaning: this step is a key step. If the data with low quality enters the model, the result is unsatisfactory. The purpose of this step is to delete duplicate data, delete or use corresponding methods to replace outliers, so as to ensure that the data is meaningful for the model establishment.

Data feature extraction: this step is often used in text analysis. For example, when analyzing the impact of stock bar comments on stock returns, comments can't be used directly. You need to extract keywords and endow them with a certain value to make the text numerical for later analysis.

Data structure: data structure means to ensure the consistency of data structure, which can greatly improve the speed of data reading and reduce the time cost in the modeling process.

Data storage: data storage is very important, for example, the banking industry can obtain natural customer information, and these data can be well preserved for a long time, which can provide the basis for future data mining. At present, MySQL and mongodb are widely used in database.

Data analysis: data analysis is the modeling process after data preprocessing. It is important to determine which model to use and the dependent variables. The commonly used models are decision tree model, neural network model, logistic regression model, clustering model and so on.

Analysis result evaluation: this step is the core of the whole data mining process, and it is the interpretation of the model results. A good model result is not equal to a good strategy. How to explain the model results in an easy to understand way and use them for decision makers is a really good model [6].

2.3 Blended Learning
In today's society, with the rise of information technology, more and more people realize the convenience brought by the combination of learning and network. "The so-called blended learning is to combine the advantages of traditional face-to-face learning with the advantages of distance network learning. In other words, it should not only reflect the role of teachers in guiding, stimulating and supervising in the teaching process, but also fully show the students' role The autonomy, participation and self creation of the subject. "This is the understanding given by Professor He Kekang, a famous scholar of blended learning theory in China. This combination of online learning and face-to-face learning not only emphasizes the situation that "students" should be the main part of learning and "teachers" should be the auxiliary part, but also stimulates students' interest in learning with the help of new media such as information network [7].

2.4 SPOC Teaching
First of all, for the lack of English class hours in higher vocational colleges, SPOC teaching mode will arrange teaching activities reasonably and decompose them effectively. Before the class, we can master the knowledge by self-study, solve the problems in the class, interact with teachers and students, and use review after class to consolidate and strengthen, and teacher guidance and teaching evaluation run through the whole process. Secondly, the traditional task-based teaching model can not achieve individualized teaching, reflecting personalized teaching [8]. SPOC teaching mode can provide high-quality teaching resources, and the design of each teaching link highlights the dominant position of learners and realizes personalized customized learning. Finally, teachers play the role of mentors to guide students to complete the learning process.

2.5 Related Formulas
Pearson correlation analysis [10]:
The sample correlation coefficient is as follows

\[ \rho = \frac{\text{cov}(X,Y)}{\sqrt{\text{var}(X)\text{var}(Y)}} \]  

(1)

Multiple linear regression theory

\[ Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \mu \]  

(3)

3. Teaching Practice Design

3.1 Analysis of Teaching Needs

The foreign language teaching and research section has done a written survey in the whole school before offering public elective courses, so as to understand the students' English learning needs. This survey uses "questionnaire star", through English teachers in the class (basically covering the whole school) to do publicity and promotion. Through the analysis of the valid questionnaires collected, it is found that the main needs of students for English learning are reflected in the following aspects: Workplace English; life English; interests and hobbies; English proficiency test and upgrading from college to undergraduate.

3.2 Analysis of Teaching Objects

The experimental class is "Advanced English class 1" and the control class is "Advanced English class 2". There are 30 students in each class, a total of 60 students. Among them, the experimental class adopts SPOC teaching mode based on blended learning theory.

According to the requirements of SPOC teaching for teaching tools and students, questionnaire survey was used to understand the students in the experimental class. 30 students in the experimental class were given questionnaires about "survey of students' network situation" and "recognition of online learning", and 30 valid questionnaires were collected, with a recovery rate of 100%.

4. Analysis of Survey Results

4.1 Overall Reliability Analysis of Learning Attitude Questionnaire

| Table 1. Overall Reliability Analysis of Learning Attitude Questionnaire |
|---------------------------------------------------------------|
| dimension | classification | number of entry | Cronbach's Alpha |
|----------|----------------|-----------------|------------------|
| Cognitive Dimensions | Evaluation of teachers | 6 | 0.893 |
| | Evaluation of teaching materials | 3 | |
| | Evaluation of the environment and resources | 5 | |
| Emotional dimension | learning interest | 6 | 0.813 |
| | Feeling in English-speaking Countries | 4 | |
| Behavioral Intention Dimension | Learning Initiative | 6 | 0.846 |
| | Focus on learning value | 3 | |
| | Willingness to overcome difficulties | 3 | |
| | The desire for interaction | 3 | |

In this pretest, 60 subjects were divided into high and low groups according to 27% of the item scores. There were significant differences in all 39 items between the high and low groups, indicating that the questionnaire had a good discrimination. The Cronbach's alpha coefficients in three dimensions are 0.893, 0.813 and 0.846, ranging from 0.8 to 0.9, indicating that the overall reliability of the
questionnaire is high, and there is a high internal consistency among the factors.

4.2 Students' recognition of Online Learning

![Figure 1. Students' recognition of online learning](image)

It can be seen from the figure above that most of the students are very much looking forward to and recognized the online learning, and fully meet the conditions for the experiment.

4.3 Post Test Analysis of English Scores of Experimental Class and Control Class

| Class       | N  | mean value | standard deviation | Standard error of mean |
|-------------|----|------------|--------------------|------------------------|
| English test scores | 1.00 | 30 | 62.7333 | 8.01694 | 1.46369 |
|             | 2.00 | 30 | 58.1667 | 9.38481 | 1.71342 |

In the table above, we can see that the average scores of advanced English class 1 (experimental class) and advanced English class 2 (control class) are 62.7 and 58.1 respectively. In the above table, we can see that the sig. Value is 0.238, greater than 0.05, which means that the variance of the two groups of data (the results of the experimental class and the control class) is equal. The sig (bilateral) value is 0.047, less than 0.05. According to this significance level, we can know that there is a significant difference between the two classes. The average score of the experimental class is better than that of the control class. Therefore, compared with the traditional English teaching, the SPOC teaching mode based on blended learning theory can improve students' overall English level and learning efficiency.

In terms of average score, the experimental class is higher than the control class, and the number of students in the three grades of "excellent", "good" and "pass" is more, and the number of students in the grade of "need to work hard" is less.

| Class       | F    | Sig. | t    | df | Sig. | Mean value difference | Standard deviation error | 95% confidence interval for difference |
|-------------|------|------|------|----|------|-----------------------|-------------------------|-------------------------------------|
| English test scores | Assumption of equal variance | 1.419 | .238 | 2.026 | 58 | .047 | 4.56667 | 2.25349 | .05582 | 9.07751 |
|             | Variance of variance          | 2.026 | 56.58 | 58 | .047 | 4.56667 | 2.5349 | .05348 | 9.07985 |
5. Conclusion

Based on the diversity of English Teaching in higher vocational colleges, combined with the “screening” characteristics of public elective courses, this paper constructs a SPOC teaching mode of public elective courses in Higher Vocational Colleges Based on blended learning theory, and carries out a small-scale teaching practice of this mode. First of all, the model effectively improves the students' learning autonomy, makes the students understand the significance of learning more deeply in the autonomous learning after class, and at the same time, actively asks questions on what they have learned before class, so as to cultivate their innovative thinking ability. Secondly, the model makes effective use of the time before, after and in class, so as to solve the problem of insufficient class hours in Higher Vocational English teaching to a certain extent. At the same time, using SPOC platform can also complete the homework and test correction, student learning data statistics, which greatly saves the energy of teachers, so that teachers have more time for personalized teaching, individualized teaching. Finally, according to the results of SPSS score analysis and questionnaire analysis, it is proved that the implementation of SPOC teaching mode based on blended learning theory is effective and successful.

References

[1] Ige O A , Hlalele D J . Effects of computer-aided and blended teaching strategies on students' achievement in civic education concepts in mountain learning ecologies[J]. Education and Information Technologies, 2017, 22(6):1-17.
[2] Kaplan A M , Haenlein M . Higher education and the digital revolution: About MOOCs, SPOCs, social media, and the Cookie Monster[J]. Business Horizons, 2016, 59(4):441-450.
[3] Zhang E , Zhang W , Jin C . SPOC-based Flipped Classroom of College English: Construction of an Effective Learning Model[J]. Technology Enhanced Foreign Language Education, 2018, 13(1):37-38.
[4] Yu H . Promoting strategy research on blended teaching mode reform in colleges and universities: a case study in China[J]. International Journal of Social Media and Interactive Learning Environments, 2017, 5(1):79-80.
[5] Caravias V. Literature Review in Conceptions and Approaches to Teaching using Blended Learning [J]. International Journal of Innovation in the Digital Economy, 2015, 6(3):46-73.
[6] Caravias V. Literature Review in Conceptions and Approaches to Teaching using Blended Learning [J]. International Journal of Innovation in the Digital Economy, 2015, 6(3):46-73.
[7] Mchunu S P , Imenda S N. The Effects of Traditional, Outcomes Based Education (OBE) and Blended Teaching Approaches in Alleviating Conceptual Difficulties and Alternative Conceptions in Grade Twelve Mechanics [J]. International Journal of Educational Sciences, 2015, 8(2):333-343.
[8] Kaplan A M, Haenlein M. Higher education and the digital revolution: About MOOCs, SPOCs, social media, and the Cookie Monster[J]. Business Horizons, 2016, 59(4):441-450.
[9] Zhang Y, Rataj K, Simpson G G, et al. Crystal Structure of the SPOC Domain of the Arabidopsis Flowering Regulator FPA[J]. PLoS ONE, 2016, 11(8),1321-1354.
[10] Wang Y. A Study on College English High-efficiency Class Based on Blended Teaching Mode of Flipped Classroom [J]. Theory and Practice in Language Studies, 2020, 10(9):1066-1067.