

Retraction

Retracted: Data Mining Technology-Based English Listening Prediction Strategy and Its Training Approach

Security and Communication Networks

Received 18 November 2022; Accepted 18 November 2022; Published 5 December 2022

Copyright © 2022 Security and Communication Networks. 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.

Security and Communication Networks has retracted the article titled “Data Mining Technology-Based English Listening Prediction Strategy and Its Training Approach” [1] due to concerns that the peer review process has been compromised.

Following an investigation conducted by the Hindawi Research Integrity team [2], significant concerns were identified with the peer reviewers assigned to this article; the investigation has concluded that the peer review process was compromised. We therefore can no longer trust the peer review process, and the article is being retracted with the agreement of the Editorial Board.

References

- [1] Y. Zhu, “Data Mining Technology-Based English Listening Prediction Strategy and Its Training Approach,” *Security and Communication Networks*, vol. 2022, Article ID 4852792, 8 pages, 2022.
- [2] L. Ferguson, “Advancing Research Integrity Collaboratively and with Vigour,” 2022, <https://www.hindawi.com/post/advancing-research-integrity-collaboratively-and-vigour/>.

Research Article

Data Mining Technology-Based English Listening Prediction Strategy and Its Training Approach

Ye Zhu 

Guang'an Vocational & Technical College, Guang'an 638000, China

Correspondence should be addressed to Ye Zhu; 171847062@masu.edu.cn

Received 20 January 2022; Revised 5 February 2022; Accepted 14 February 2022; Published 7 March 2022

Academic Editor: Muhammad Arif

Copyright © 2022 Ye Zhu. 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.

The big data era of “data-driven schools, analysis, and change education” has arrived, and the technology of data mining was born in the education industry. Based on data mining technology, this study explores English listening prediction strategies and training approaches. Prediction is an effective learning strategy in listening comprehension. Cultivating students to use predictive strategies is helpful to improve their listening comprehension, mining and analyzing the listening data generated by English skill training system, selecting the data related to students' listening as features, aiming at students' test listening scores, carrying out model training, and realizing the prediction of students' listening. After evaluation and analysis, it is found that data mining technology can be used to predict students' listening more accurately.

1. Introduction

The Internet revolution has promoted the development of distributed databases and server systems. There are huge and complicated data information systems in all fields of the world [1]. How to mine rules and patterns from the existing information base in time becomes the primary task of system analysts. In recent years, education reform has been one of the key issues of social concern. With the continuous progress of society, high-end science and technology and products have been widely used in all fields of society, which makes people's quality of life constantly improve [2]. Similarly, the teaching quality of the education industry is constantly improving, and the teaching methods, means, and teaching environment are in the process of upgrading. Prediction is a basic strategy to use language knowledge and original knowledge in mind to understand information [3]. An important microskill that students need to master in English listening is prediction. Prediction is a basic strategy to use language knowledge and original knowledge in mind to understand information. With the rapid development of the Internet, big data comes into being, which makes data mining technology constantly applied in the field of education, providing convenient teaching conditions for

schools, teachers, and students, and the application of data mining technology in the field of education has also received extensive attention [4]. As a receptive skill, listening is one of the four basic skills in foreign language learning. Listening is an important part of language teaching. Many studies at home and abroad have found that the process of listening comprehension is closely related to the application of listening strategies [5]. There are obvious differences in the use of listening strategies between good and poor learners. The former can flexibly use various learning strategies to help them complete various listening tests. Therefore, listening teaching should be guided by learning strategies.

With the rapid development of China's economy, China's requirements for personnel training specifications are getting higher and higher, while the old educational ideas, educational objectives, and educational methods are far from meeting this requirement [6, 7]. Therefore, as an advanced skill in listening, the prediction has received extensive attention. If students can understand the skills related to prediction and apply them consciously so that their prediction ability can be continuously improved, their initiative in English listening can be greatly increased [8]. Listening comprehension is an abstract thinking activity, and listening comprehension is a kind of rapid

comprehensive responsibility [9]. It includes the ability to distinguish sounds, the ability to identify various grammars and sentence patterns, the ability to predict, the ability to think correctly, and the ability to combine knowledge with experience, etc. [10]. In the study of learning strategies, listening strategies have aroused great interest of scholars. If listening prediction is effectively used in the three stages of listening activities (before, during, and after listening), it will help students improve their listening situation, cultivate good prediction habits, and enhance their confidence in listening ability. Based on the research of data mining technology, this paper puts forward English listening prediction strategies and training ways.

The process of regionalization and globalization of the world economy is accelerating, and the political, economic and cultural exchanges among countries are deepening day by day [11]. Especially today's rapid development of science and technology has greatly aggravated this contradiction. In international exchanges and cooperation, in academic exchanges, when learning advanced technology and experience from other countries, we are no longer satisfied with written exchanges, but tend to face-to-face and direct exchanges and cooperation [12]. This requires our talents at all levels to have good listening and speaking skills. Language communicative competence can be divided into listening, speaking, writing, and reading. The ability of listening comprehension and prediction is one of the most powerful auxiliary factors of language acceptance [13]. When people listen to others, they often use their concepts and social knowledge of the main contents of the upper and lower parts to limit the scope of the words they are about to hear and predict the next content. English listening comprehension skills mainly include the following abilities: the ability to predict and choose, the ability to recognize sounds by listening, the ability to screen important information, the ability to capture keywords, the ability to analyze and summarize the main idea, the ability to speculate the intrinsic meaning, and the ability to remember coherently [14]. Among them, predictive ability occupies a very important position. In English listening prediction, there are many factors that affect listening prediction, and it is necessary to make a comprehensive analysis of each factor. In the era of big data, how to find valuable information from a large amount of data and use this information to predict unknown or future values is becoming more and more important. Data mining technology explores the correlation between information by building correlation models.

2. Related Work

Literature [15] uses data mining technology to simulate and predict English scores. Literature [16] analyzes the scores of CET-4 based on data mining technology. Literature [17] holds that data preprocessing is the data preparation work before data mining, and the quality of data is the prerequisite for the prediction results. Literature [18] holds that when learners hear the first half of the listening material, they can guess the exact meaning of the following text according to what they have heard, even if they do not guess the exact

words. This is "prediction" in listening. Literature [2] holds that prediction is one of the important strategies that must be cultivated in listening class teaching. If the prediction strategy is combined with the teaching content, it can effectively promote learners' learning. Literature [19] encourages teachers to "predict according to the communicative context" as an effective strategy when training students' listening strategies. Literature [20] designs the traffic flow prediction system based on data mining technology. Literature [21] studies the concept and common methods of data mining. Literature [22] Using data mining technology, the factors affecting students' achievement are analyzed more comprehensively, and then targeted measures are taken to further improve the teaching quality.

3. Methodology

3.1. Data Mining. From the application point of view, data mining, as a brand-new data analysis technology, is mainly for business objects [23]: Extracting target data from massive commercial information, cleaning and processing it, and finally mining and analyzing the target data, the process of discovering patterns and laws of commercial value that people did not know before [24]. Applying the data mining method to the field of education can mine and analyze a large amount of data from the education system and obtain valuable information and knowledge contained therein, which can serve not only teachers, students, and researchers in the field of education but also system managers and software developers [25].

Data mining starts with the original data from the real world, models the data through specific algorithms, and finally obtains the knowledge that people are interested in, and then applies this knowledge to practice to guide the work [26]. There are many technologies of data mining, and there are different classifications according to different classifications. Before data mining, it is a necessary process to explore the original data foreseeably, which can avoid drawing wrong conclusions due to blind analysis. Usually, technicians who conduct data mining need to know the background information of the data in advance and have in-depth communication with the demanders. According to the type of target data and the task of data mining, they finally determine one or several algorithms for data mining. Data mining deals with massive data, which is generally stored in the database system and is the result of long-term accumulation. However, it is often not suitable for direct mining, and data preprocessing is needed. Whether this preparation is sufficient or not has a key impact on the efficiency and correctness of the mining algorithm. The data mining process of English listening prediction is shown in Figure 1.

Data mining can generally be divided into four stages: problem definition, data preparation, data mining, evaluation, and presentation. Each of these four stages has certain requirements. If a certain stage fails to achieve the expected goal, it is necessary to stop the current process, go back to the previous step to make adjustments, and then implement it again. Generally speaking, data mining is a process in which all steps are interrelated and cyclic. The information finally

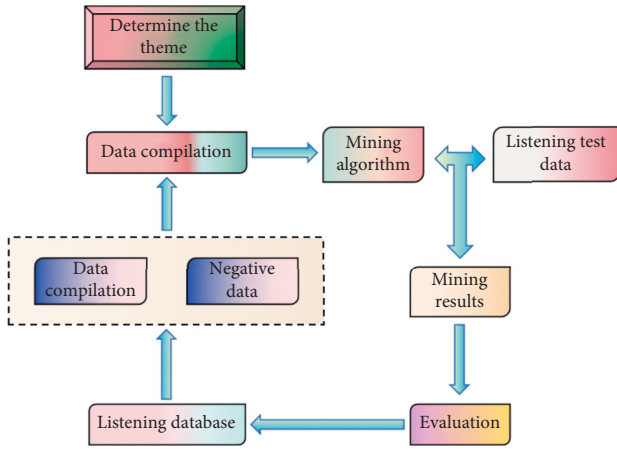


FIGURE 1: Data mining process of English listening prediction.

mined by data mining may not necessarily meet people's requirements for results, so it is necessary to evaluate the effectiveness of data mining results to determine whether the results can be used. Because the final results of the analysis are presented to users, we should show the mining results as easily as possible.

Common methods of data mining include statistical analysis method, decision tree method, support vector machine, neural network, genetic algorithm, and so on. A decision tree is a tree-like model, which can be regarded as the combination of rules of if-then. This algorithm is often used in classification and prediction problems. At this stage, the most critical operation is to choose the best partition method on the nodes of the tree. The decision model of English listening prediction based on data mining is shown in Figure 2.

Classification and prediction are two different forms of data analysis in data mining, which can be used to extract data classes that people are interested in and can also be used to predict the development trend of future data. This is an instructive learning process. Classification algorithms can be divided into two categories according to the model. The predictive model will directly tell you which category the data should be classified into, while the probabilistic model will not directly tell you the result but will tell you the probability of belonging to each category. A prediction algorithm refers to the use of historical data to predict future development trends. The commonly used prediction models are regression analysis and time series prediction.

3.2. English Listening Prediction Strategy and Its Training Approach. Prediction in listening mainly refers to the fact that students do not passively receive the information they are listening to but use an active learning attitude to predict what they are listening to with their own knowledge and information. Predictive microskills are mainly based on the research results of psycholinguistics.

The accuracy of prediction refers to the gap between the content predicted by the listener and the following text of the speaker. The smaller the gap, the higher the accuracy. Before listening or when listening to the above, you can know the

exact words below, and we say that this prediction has the greatest accuracy. The purpose of English listening teaching is to enable students to have the ability to successfully understand real phonetic materials. There are great differences among English learners in listening comprehension. However, there are several reasons for poor learners' hearing impairment. ① Some English learners are unable to adapt to the speed of abbreviation, linking, and change in spoken English due to a lack of phonetic knowledge. So as to be unable to respond quickly to the listening content, which causes the obstacle of listening comprehension. ② Lack of knowledge of western cultural background and obstacles to understanding many culture-loaded words also increase the difficulty of listening comprehension. ③ Psychological factors, including nervousness and anxiety also caused obstacles in English listening.

Before mining the collected data, it is necessary to have a general understanding of the basic information of the data and be aware of the data to be studied. The purpose of analyzing the prediction process and its characteristics is to find out the factors that affect the prediction process theoretically. We think that the factors that affect the prediction process include at least the following prediction motives, clues, working memory processing ability, relevant knowledge and experience, and speech output ability. The process of data mining is actually a process of deepening the multidimensional data view. Because different types of data tables from multiple databases are collected, it cannot meet the needs of data mining. In order to improve the efficiency of data mining and increase the reliability of mining results, we should preprocess the original data before starting data mining.

In the process of decision tree construction, the node with the highest information gain will be selected as the current split node, which will make the information required by the finally generated tree structure to divide the training samples minimum. Assuming that the training data set S is divided according to the category attribute C , its classification information entropy is as follows:

$$H(S, C) = - \sum_{i=1}^m p_i \log_2(p_i), \quad (1)$$

where m represents the total number of classes, and p_i represents the probability that the i th class appears in the entire training tuple. Assuming that the training data set S is divided according to the conditional attribute A , the classification information entropy of the conditional attribute A dividing S relative to C is as follows:

$$H(S, A|C) = - \sum_{j=1}^v \frac{|S_j|}{|S|} H(S_j, C). \quad (2)$$

The information gain of attribute A splitting dataset S is denoted as follows:

$$\text{gain}(S, A|C) = H(S, C) - H(S, A|C). \quad (3)$$

In native language listening, the ability to predict seems to be an instinctive response that makes people feel at ease

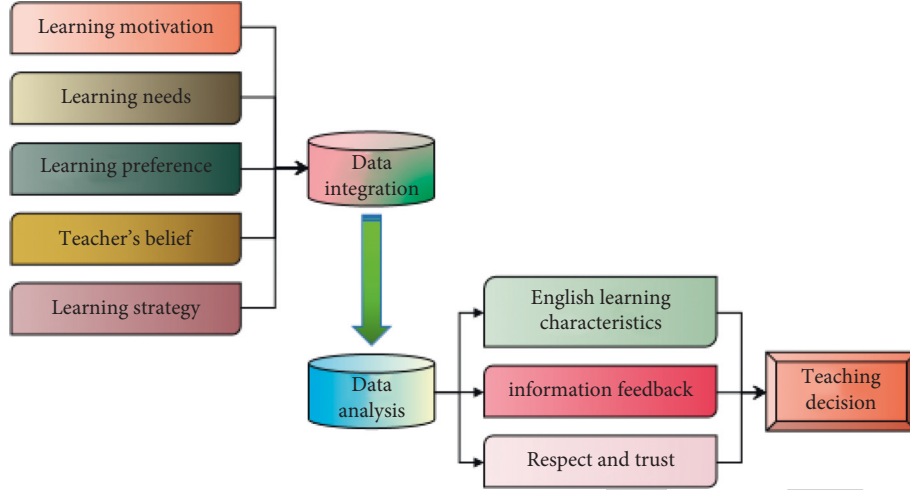


FIGURE 2: English listening prediction decision model based on data mining.

when listening and comprehending their native language. Listeners often predict what the speaker will say in the next sentence, and sometimes it is even accurate. Even if the prediction is wrong, they can adjust immediately to keep up with the information input. However, in English listening comprehension, this seemingly instinctive response is very weak, which causes the phenomenon of English listening comprehension difficult. Due to the limited time for foreign language learning and the opportunities to contact and use foreign languages, the formation of foreign language prediction skills must have its own characteristics. Therefore, in order to improve English listening, we should try to cultivate and train this predictive ability so that it can be gradually integrated into ordinary English listening and become an instinctive response like a mother tongue.

In the application, the data is described as a relationship, based on the fact table and the dimension table, and the aggregate function is used to realize the data analysis. Prediction helps listeners receive listening material with purpose. In the listening test, the listening content can be predicted according to the questions and options before listening so that the content of the material can be predicted as needed during the listening process, and the content of the material can be selectively accepted, carefully listened to, and unimportant ones ignored or skipped. The plot helps to focus attention on key content. The data structure processed by the multidimensional data model is a multidimensional data structure, and the distribution of data in the multidimensional space is always sparse and uneven. At the location of the event, the data is aggregated, and its density is high. Integrate data, combine data from different databases into one data table. Here, the basic information and English listening score information are combined with student number and name as keywords, and duplicate attributes and attributes that are redundant or have little relationship with the mining task are deleted. Eventually, a new data table was formed. Calculate the total information entropy required to classify a given sample according to formula (1).

$$(x_{LT}^i, y_{LT}^i) = \sum_{j=1}^k w_j (x_{RT_j}^i, y_{RT_j}^i). \quad (4)$$

The information entropy of each attribute is calculated according to formula (2).

$$W_j = \frac{d_j}{\sum_{j=1}^m d_j}. \quad (5)$$

The attribute splitting information metric is calculated according to formula (3).

$$HWt = \frac{\sum_{i=1}^N D_i(x)}{N}. \quad (6)$$

The online processing of composite mode is that some summary data are stored in memory, while basic data and other summary data are stored in a relational database. If the lower level is relational, the upper level is a multidimensional matrix. This method has better flexibility. Speech speed and pause are two adjusting variables of this test. The faster you speak, the more stimulation per unit time, the more information your brain needs to process, and the greater the load. The so-called speed of speech is actually the listener's feeling. When his processing speed can catch up with the speed of speech, he will feel unhappy, but when he cannot catch up, he will feel fast. Therefore, the speed of speech is limited by the speed of speech processing of the listener. In listening comprehension, if the existing schema knowledge of phonetics, vocabulary, grammar, and cultural background in the listener's brain is activated, it can be the basis for predicting the following text. Pause is another variable. Before the listener chooses the answer, he needs to connect the information he hears with the text information and make an inference. If his inference time exceeds the pause time, he will feel that the pause is short. In short, the degree of response to these two variables represents the listener's speech processing ability, which is his listening level. In this test, these two variables have a significant effect on the listener. In the classification problem, assuming that there are q classes,

and the probability that the sample points belong to the J class is m_{jq} , the index of the probability distribution is defined as follows:

$$ES_i = \sum_j \left(1 - \sum_q P_{iq} m_{jq} \right), q \neq i, j. \quad (7)$$

If the probability that the sample point belongs to the first class is AE_i , then the exponent of the probability distribution is as follows:

$$AE_i = \frac{ES_i}{S_i} = \frac{\sum_j (1 - \sum_q P_{iq} m_{jq})}{\sum_j}. \quad (8)$$

Here, S_i is the subset of samples belonging to the q th class in E , and i is the number of classes.

Due to the few features that can be obtained, and some features that have collinearity problems, traditional regression analysis and other methods are often not very effective. This is also one of the main reasons for using a data mining algorithm in this paper. In a multicube structure, a large data structure is divided into multiple multidimensional structures. These multidimensional structures are a subset of the dimensions of big data, and the dimensions are partitioned for a specific application. That is, the hypercube structure becomes a subcube structure. It has strong flexibility and improves the efficiency of data analysis. Generally speaking, the multicube structure is more flexible. If the analyst has good multidimensional analysis capabilities, it is better to use the hypercube structure because it has good view flexibility and flexibility.

4. Result Analysis and Discussion

The data mining stage is to mine and analyze the processed data, which is the core stage of the whole data mining and needs the support of the experience of related industries. We need to choose the appropriate mining algorithm for analysis according to the mining tasks and objectives determined in the problem definition stage, which can be clustering, classification, prediction or association rule mining, etc. Prelisting activities activate the background knowledge related to listening materials in the listener's long-term memory so that the listener can use this background knowledge to promote his understanding after perceiving the auditory information. In addition, because the process of foreign language listening comprehension is the same as that of native language listening comprehension, prelistening activities can enhance students' awareness of prediction, thus applying unconscious prediction in the native language listening environment to the target language listening environment. Due to the complexity of the training of the prediction process and the short research process, it is difficult for us to test all the hypotheses. We can only determine the appropriate goals according to the subjects and selectively carry out teaching practice.

One of the experimental classes implemented the traditional forecasting method, which is called the traditional group. The other class adopts the new method, which is

called the new method group. Almost no students have had prediction experience in foreign language listening activities, and there is no significant difference in previous foreign language scores. In order to confirm that the listening level of the three classes is equal, we conducted three tests before the experiment. The test results are shown in Figure 3. The results show that there is no significant difference in the scores of the three classes.

Exploratory data analysis is a process of analyzing data sets to decide which method to choose for statistical inference, also known as descriptive statistical analysis. Different from the traditional statistical analysis methods, which need to meet certain assumptions in advance, exploratory data analysis is very flexible in dealing with data objects in an incomplete formal way. Hypercube refers to describing an object with three or more dimensions, and each dimension is perpendicular to the other. The measured value of data occurs at the intersection of dimensions, and all parts of the data space have the same dimension attributes. There is a deformation of the hypercube structure, that is, the contraction hypercube structure. This structure has higher data density and fewer data dimensions and can add additional analysis dimensions. Generally speaking, end users are easy to approach the hypercube structure, which can provide high-level reports and multidimensional views. Ask students to repeatedly use the method of obtaining the same clues, use clues in different situations, and ask students to grasp these clues according to different situations. At first, students can find clues and reach a certain speed, but when it comes to comprehensive application, it should be carried out simultaneously with other processes and completed in an instant.

When there are returned samples, the weight of those observations with a high misjudgment rate will be increased so that there may be more observations in the new samples that were wrongly divided before. Then a new classifier will be formed to enter the next iteration. In each iteration, the misjudgment rate will be given to the classifiers generated in this round. The final result will be produced by the classifiers in each stage by weighted voting according to the misjudgment rate. The data collection methods are tests, questionnaires, and classroom observation. In the experiment, two predictive level tests and two related surveys were conducted, respectively, and the listening scores of the midterm and final exams were collected as data. The tests were arranged in the 10th week after the experiment and at the end of the experiment, and a questionnaire survey was conducted at the same time. The results were analyzed. The listening scores of the midterm exam are shown in Figure 4. The listening scores of the final exam are shown in Figure 5.

The ability of clue integration affects the sufficiency of clues. Integration does not mean simply listing all the clues but also knowing their inner and deep relationships so as to discover the context and trend of things development and make a correct prediction.

Most of the traditional statistical prediction models have certain requirements or assumptions for data, and the model itself has a clear mathematical form. However, in most cases, we cannot make any assumptions about the distribution of

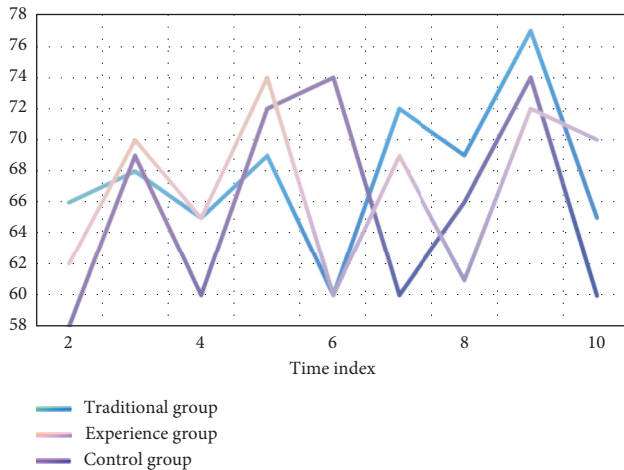


FIGURE 3: Students' scores before the test.

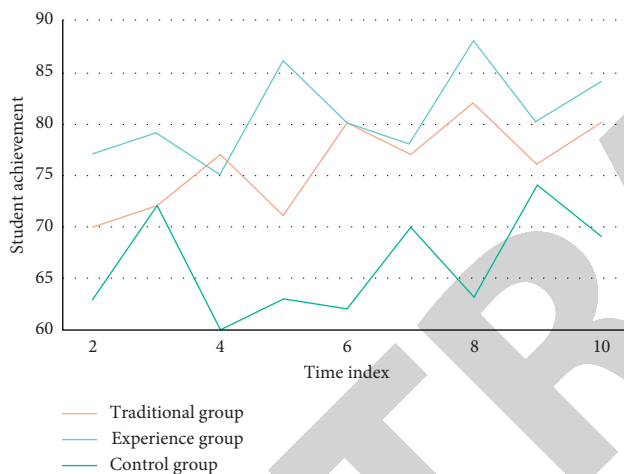


FIGURE 4: Listening scores of midterm exams.

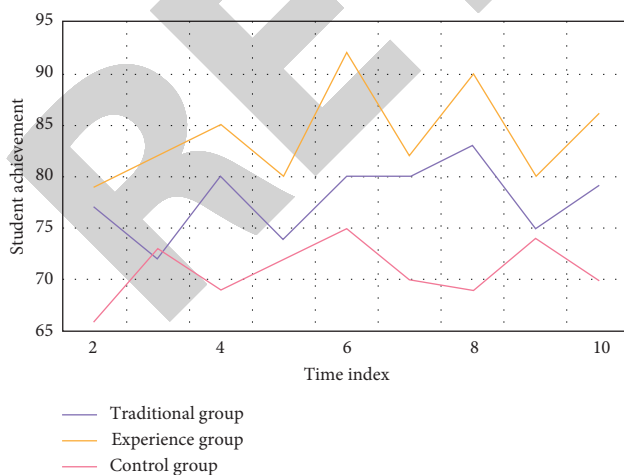


FIGURE 5: Listening scores in the final exam.

real data. In addition, limited mathematical formulas cannot describe all the situations in the real world. At this time, the machine learning method shows its unique advantages. It

does not need any restrictions on data, and it is often more accurate than the classical methods. Whether we can make a prediction or not, the sufficiency of clues is the key. The sufficiency of clues means that the clues provided when the above occurs have strong restrictions on the following. That is, a strong context is formed. Only with sufficient clues can relevant knowledge and experience be activated and restricted so that prediction can happen, so sufficient clues are a necessary condition for prediction to happen. Through data analysis, we study the enthusiasm of students and get the results as shown in Figure 6.

With the increase of forecasting practice, more and more students' forecasting ability is improving, and the new method group is obviously higher than the traditional group. The statistics of "making prediction exercises interesting" have the same result. Therefore, we can conclude that the new method can make students have a positive attitude towards prediction more than the traditional method. When students get a sense of success from forecasting, their interest will be enhanced, which in turn will promote the development of forecasting ability, and the role of forecasting can be better played. Listening training generally includes three stages: prelistening, listening, and postlistening training. Prelistening training is the preparation stage. Organizing prelistening activities by listening to teachers in the preparation stage helps students to use prediction strategies more effectively. Select the average misjudgment rate of the test set in 10% cross-validation to judge whether the model is good or bad. In the process of 10-fold cross-validation, 14 pieces of data are taken out for training each time, and the remaining one piece is tested to calculate the misjudgment rate of the test set. This cycle is repeated for 20 times, and finally, the calculated misjudgment rate of 20 times is averaged. From general category association to special topic association, from simple to complex, from slow to fast, from control to freedom, from voice to silence, each association should be repeated many times until it is proficient, and students are encouraged to develop the habit of intentional or unintentional association. To prove the practicability of this method, we are testing. When testing with Test1, students are not given any time to read before, and they start listening as soon as they get the test paper. Before the Test2 test, students have 3 minutes to read ahead, but they do not give any guidance. Before the Test3 test, instruct students to predict each long conversation, essay, and news according to the content of questions and options in detail. This guidance lasted about 20 minutes. Results: There were significant differences in the scores of the three tests as shown in Figure 7.

When evaluating the accuracy of the decision tree, we mainly measure it from the confusion matrix. The confusion matrix is an effective method to identify tuples of different categories by analyzing classification decision trees. In the confusion matrix, rows are real classes, and columns are prediction classes. The self-help method puts back sampling, each training set is independent, and the weight of each sample set is equal. This algorithm has no restrictions on the base learner during integration, and the most commonly used ones are decision trees and neural networks. In the

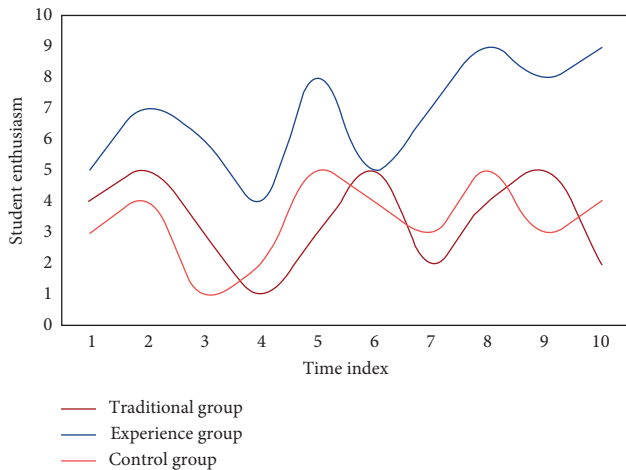


FIGURE 6: Comparative results of students' enthusiasm.

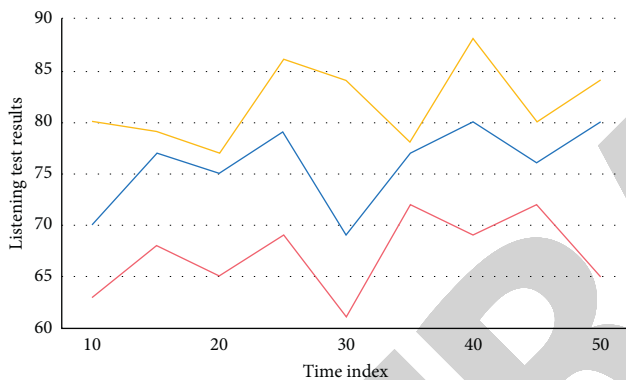


FIGURE 7: Results of three listening tests.

experiment, teachers are required to supervise students' self-practice, collect feedback information at any time and give timely guidance and encouragement.

5. Conclusions

With the advent of the era of big data, data mining is infiltrating into various industries, and data mining technology is changing our lives. Reasonable use of data mining technology can not only improve efficiency but also effectively avoid risks. Through big data, taking English listening teaching as the research object, applying data mining technology, this paper analyzes the teaching environment, teachers' strength, teaching process, students' learning methods, learning motivation, learning time, and other factors, and makes statistics on the effect of English listening teaching. This paper analyzes the problems and conditions faced by the cultivation of prediction skills, puts forward the objectives and methods of the cultivation of listening prediction skills, and makes a teaching attempt. Students should fully realize the importance of prediction in three stages: before listening, during listening, and after listening. Forecasts point to the future. We should analyze the discourse and make some inferences in our understanding. On the one hand, we can deduce the implied relationship

between propositions, the intention of speakers, and their emotional state from what we have said. On the other hand, we can infer the possibility of the intention and emotion of the speaker or the topic figure that will appear below, based on the known information obtained in front, common sense, and experience, so as to predict and look forward to it. From the test scores and the results of the questionnaire survey, we can see that cultivating prediction skills is beneficial to the improvement of students' listening scores. And the effect of the new method is obviously greater than that of the traditional method, and the effect of the new method in cultivating prediction skills is obviously better than that of the traditional method.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The author declares no conflicts of interest.

References

- [1] S. Lin, "Data mining artificial intelligence technology for college English test framework and performance analysis system," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 2, pp. 1–11, 2020.
- [2] G.-J. Hwang, T.-C. Hsu, C.-L. Lai, and C.-J. Hsueh, "Interaction of problem-based gaming and learning anxiety in language students' English listening performance and progressive behavioral patterns," *Computers & Education*, vol. 106, no. 3, pp. 26–42, 2017.
- [3] Y.-C. Kuo, H.-C. Chu, and M.-C. Tsai, "Effects of an integrated physiological signal-based attention-promoting and English listening system on students' learning performance and behavioral patterns," *Computers in Human Behavior*, vol. 75, no. 10, pp. 218–227, 2017.
- [4] M. F. Scudder, "Beyond empathy: strategies and ideals of democratic deliberation," *Polity*, vol. 48, no. 4, pp. 1–27, 2016.
- [5] N. Nga, "Understanding the impact of listening strategy instruction on listening strategy use from a socio-cultural perspective - ScienceDirect," *System*, vol. 81, pp. 63–77, 2019.
- [6] J. Liu, G. Shi, J. Zhou, and Q. Yao, "Prediction of college students' psychological crisis based on data mining," *Mobile Information Systems*, vol. 2021, no. 23, 7 pages, Article ID 9979770, 2021.
- [7] P. Wingo, "How suitable are TED talks for academic listening?" *Journal of English for Academic Purposes*, vol. 30, pp. 79–95, 2017.
- [8] L. Shidong, T.-H. Kuo, M. ul Islam, M. Talha, S. A. Lone, and S.-Y. Wei, "The impact of sports activities on economic survival, social life of women, and what do big data analysis reveals about social life of sports women," *Revista de Psicologia del Deporte*, vol. 30, no. 3, pp. 229–241, 2021.
- [9] K. L. B. Deroey, "The representativeness of lecture listening coursebooks: language, lectures, research-informedness," *Journal of English for Academic Purposes*, vol. 34, pp. 57–67, 2017.
- [10] Yu.-L. Cheng, "Unfamiliar accented English negatively affects EFL listening comprehension: it helps to be a more able accent

- mimic," *Journal of Psycholinguistic Research*, no. 3, pp. 1–13, 2018.
- [11] K. L. B. Deroey, "The representativeness of lecture listening coursebooks: language, lecture authenticity, research-informedness," *Journal of English for Academic Purposes*, vol. 34, pp. 57–67, 2018.
- [12] A. Becker, "L2 students' performance on listening comprehension items targeting local and global information," *Journal of English for Academic Purposes*, vol. 24, pp. 1–13, 2016.
- [13] A. Y. W. Chan, "How much do Cantonese ESL learners know about the English article system?" *System*, vol. 56, pp. 66–77, 2016.
- [14] S. Cao, R. Chen, H. Liu, and R. Shi, "An empirical study on multimodal discourse analysis of college English teaching in the context of new media," *Journal of Intelligent and Fuzzy Systems*, vol. 2021, no. 2, pp. 1–5, 2021.
- [15] N. Chu and W. Ma, "Distribution of large-scale English test scores based on data mining," *Complexity*, vol. 2021, no. 19, 10 pages, Article ID 5531595, 2021.
- [16] X. Jie and L. Yang, "CET-4 score analysis based on data mining technology," *Cluster Computing*, vol. 22, no. 5, pp. 1–11, 2018.
- [17] H. Wang, "Analysis and prediction of CET4 scores based on data mining algorithm," *Complexity*, vol. 2021, no. 12, 11 pages, Article ID 5577868, 2021.
- [18] Y. Xin, "Analyzing the quality of business English teaching using multimedia data mining," *Mobile Information Systems*, vol. 2021, no. 12, 8 pages, Article ID 9912460, 2021.
- [19] Q. Chen and D. Ou, "The effects of classroom reverberation time and traffic noise on English listening comprehension of Chinese university students," *Applied Acoustics*, vol. 179, no. 4, Article ID 108082, 2021.
- [20] S. Wang, "Listening test training in Beijing," *Journal of the Audio Engineering Society*, vol. 67, no. 5, p. 340, 2019.
- [21] C. Angeli, S. K. Howard, J. Ma, J. Yang, and P. A. Kirschner, "Data mining in educational technology classroom research: can it make a contribution?" *Computers & Education*, vol. 113, no. 10, pp. 226–242, 2017.
- [22] F. Yang and F. W. B. Li, "Study on student performance estimation, student progress analysis, and student potential prediction based on data mining," *Computers & Education*, vol. 123, pp. 97–108, 2018.
- [23] A. Pearce, S. Isabelle, and H. Francois, "Methods of assessing the rank order of prediction models with respect to variance of listening test ratings," *Journal of the Audio Engineering Society*, vol. 67, no. 2, pp. 60–68, 2019.
- [24] F. Gao, "Establishment of college English teachers' teaching ability evaluation based on Clementine data mining," *Journal of Intelligent and Fuzzy Systems*, vol. 38, no. 6, pp. 1–9, 2020.
- [25] J. Xu and Y. Liu, "CET-4 score analysis based on data mining technology," *Cluster Computing*, vol. 22, no. 2, pp. 3583–3593, 2019.
- [26] Y. Wu, "Research on interactive model of English translation based on data mining," *International Journal for Engineering Modelling*, vol. 31, no. 1, pp. 273–279, 2018.