Journal of Information Technology Review

DLINE JOURNALS

Print ISSN: 0976-3511
Online ISSN: 0976-2922
JITR 2025: 16 (1)
https://doi.org/10.6025/jitr/2025/16/1/11-18

Prediction Study of Vocational College Students' English-Speaking Proficiency based on Decision Tree Algorithm

Fan Wang
Anyang Secondary Vocational and Technical School
Anyang, Henan, 455000, China
chenxi198491@126.com

ABSTRACT

With the continuous advancement of modern technology, various information technologies have gradually matured and been applied in education. The design of English-speaking teaching in vocational colleges has been a hot topic of concern for many researchers, and the actual situation of students learning English speaking is also a major concern for educators. Up-to-date educational methods and diverse teaching approaches have revitalised English-speaking classrooms in ordinary colleges. This paper uses the decision tree algorithm in deep learning to analyse vocational college students' English-speaking proficiency in-depth and establish a predictive system for speaking learning effectiveness. Data analysis evaluates students' English-speaking abilities and levels, measuring changes in English-speaking effectiveness and providing reliable factors for actual learning situations. The association algorithm is used to optimize the decision tree predictive performance, predicting the English-speaking proficiency of vocational college students and providing assistance for subsequent teaching. The research shows that the decision tree algorithm achieves high accuracy in the statistical and predictive analysis of vocational college students' English-speaking proficiency, which greatly promotes the optimization of teaching strategies and methods.

Keywords: Decision Tree Algorithm, Vocational College, English Speaking, Prediction of Proficiency

Received: 18 August 2024, Revised 10 October 2024, Accepted 20 October 2024

1. Introduction

The learning ability and level of students in vocational colleges are complex and not stratified, with a wide range of English foundations and abilities among students [1]. English theory and English-speaking teaching have always been the core subjects that teachers and schools focus on. Vocational colleges often face the challenge of

teachers carefully designing the teaching process, but students' learning outcomes and grades are not ideal. Therefore, it is challenging for teachers to accomplish the planned teaching design in the classroom. To improve students' overall English proficiency, besides theoretical learning, practical exercises in English speaking are essential [2]. First, from the background investigation, we analyze the purpose of English teaching and the importance of English speaking. In actual teaching classrooms, the ultimate goal of English learning is to achieve seamless conversion between two languages and express ideas fluently in daily life using English [3]. The highest level of English learning is to combine personal insights with words, expressing one self while reasonably applying contextual syntax to enhance the infectivity of the language environment. The use and learning of English speaking are the most crucial parts of English subjects, meeting our basic usage needs. Using spoken language to vividly and concisely express images enables people to engage in smooth interactive activities [4].

As we all know, learning English is a prerequisite for promoting international communication. Chinese characters are one of the most complex languages in the world, and many dialects affect students' understanding of the written language. Various language dialects can only be passed down orally, and the obstacles in communication are relatively minor. This indirectly proves that practice and application are the best way to learn a language [5]. To become proficient in English speaking, teachers should create an English communication environment and design more task-based dialogues for students. At the same time, combining computer information technology to apply informational teaching methods to display English-speaking fragments encourages students to use English for communication and interaction [6]. In addition, students are prone to be affected by vocabulary and grammar misconceptions during their studies. Simply memorizing words and grammar does not guarantee improvement in English speaking proficiency. This sequential learning mode is unsuitable for students with different learning styles. Since many people do not have a good learning method without a language environment, improving English speaking proficiency is indeed a difficult problem. Therefore, China's English teaching has created a satirical term, "mute English." Creating a good language environment and using effective teaching methods are important, but good practice activities are critical for improving English speaking proficiency. Modern technology helps English learning construct a dynamic virtual scene, allowing students to naturally convert language images in their minds and form one-to-one associative relationships. To further explore the actual changes in vocational college students' English-speaking proficiency, we use the decision tree method in the deep learning algorithm to conduct statistical and predictive analysis.

2. Decision Tree Algorithm and the Current Situation of English-Speaking Development

English learning mainly includes four aspects: speaking, writing, reading, and listening. Students in vocational colleges have a better grasp of reading and writing skills, but their performance in English speaking is not satisfactory [7]. This phenomenon is mainly due to school teachers' insufficient emphasis on English speaking and outdated teaching concepts. The poor English-speaking ability of students is related to their time allocation in daily practice. Most English- speaking training occurs in class room activities, where teachers use bilingual teaching to complete dialogue tasks with students, effectively training their speaking ability [8]. However, outside of the classroom, students hardly have time for oral practice, and extracurricular assignments are mostly focused on theoretical learning. In addition, many teachers in vocational colleges believe that reading aloud and recitation are the main ways to train students' speaking ability. The flaw in this viewpoint is that it easily hinders the natural connection and subjective initiative of English speaking. Oral learning requires a strong and flexible context to help students transform the cognitive structure of their existing grammar. Excessive reading aloud can only solidify existing thinking and cannot develop students' speaking ability. In Chinese English teaching, it

is common to find students who can seemingly read English articles fluently but still struggle to organize language and express their thoughts in actual communication. Abundant evidence has shown that English-speaking contexts are of great help in learning the English language, but vocational colleges still lack training and establishment of English-speaking contexts [9].

With the advent of the Internet and computer era, online teaching and modern teaching methods are becoming increasingly widespread in education. The most obvious feature of information technology is the use of a large amount of data in daily needs. During the extraction and processing process, these massive data have certain effects and interventions on the calculation results through their own rules. As a discrete function computing method, the decision tree has shown good results in data classification and preprocessing [10]. Its feature value optimization has a significant role in enterprise performance evaluation, redistributing the original data set to form a new one. Incorporating the optimized feature values from the analysis and processing of data subsets into enterprise evaluation can serve as a judgment of value. In addition, the decision tree algorithm also has a centralized optimization approach, treating processed data as a vast information platform and dividing its internal structure into network partitions. The data content is divided into several processing batches, with temporarily unused data stored in the structure, and only the effective parts are selected for calculation. This mode ensures fast read-write speeds in prediction and is more suitable for fields with high mano euvrability. We found that the decision tree algorithm has significant advantages in data prediction, and when dealing with the complex object of students' learning situations, adopting the decision tree algorithm can address the problem of evaluating and detecting the effectiveness of vocational college students' English-speaking proficiency.

3. English Speaking Analysis System and Performance Prediction

3.1. English Speaking Analysis System

English, as the world's main language, has become even more important in the context of economic globalization and trade. With the advancement of modernization, there have been many changes in English teaching and learning methods. Information technology, such as the Internet, has provided a more convenient and reliable platform for English learning. As the mainstream language, modern teaching methods have brought English learning closer to daily life, leading to significant innovations in how vocational college students learn English. We analyzed several key teaching points in English teaching, and English speaking is a crucial aspect where students' English proficiency undergoes qualitative changes. This article mainly focuses on studying vocational college students' English-speaking proficiency, establishing a dynamic knowledge structure supported by the decision tree algorithm, and recommending personalized learning plans based on students' average individual characteristics and cognitive abilities. The main structure of the system includes English speaking proficiency assessment, data extraction and prediction using the decision tree algorithm, and learning effectiveness evaluation.

The decision tree algorithm in the data module randomly acquires vocational college students' English-speaking scores and stores the data in the integrated system. The data content is authentic and reliable, sourced from valid information on various school academic platforms. The data is divided based on various features such as students' age, duration of learning English, and theoretical English scores. At the same time, an English-speaking test module is established, utilizing the school's information technology unified platform to allow students to undergo ability testing in a standardized manner. The scoring system will automatically generate score information based on students' performance levels, evaluating specific aspects such as English

pronunciation, stress allocation, and fluency. To accurately predict the data, we use a nonlinear correlation to reconstruct the prediction model. Before statistically analyzing the data samples, we integrate the data and analysis results collected in the earlier stages according to the time sequence and embed the data sample length and feature values into the new model. At this point, the data prediction model we construct is as follows:

$$x_p = x(t_1 + \Delta t) = h[z(t_0 + bt)] + w_n \quad (1)$$

Among them, h represents the characteristic distribution trajectory of the data sequence, and based on the calculation error of the data, it can be determined whether the effectiveness of predicting the English-speaking level of vocational college students in the early stage is above the standard accuracy coefficient. Re-group the sample features, map the characteristic changes in students' learning onto trajectories, and reconstruct the spatial trajectories as follows:

$$X = [S_1, S_2, ..., S_k] = (x_n, x_{n-m}, ..., x_{n-(m-1)}) \quad (2)$$

In the formula, X represents the feature similarity of the data in the same time series. Finally, we will present the structure diagram of the optimized decision tree English speaking situation prediction model as follows:

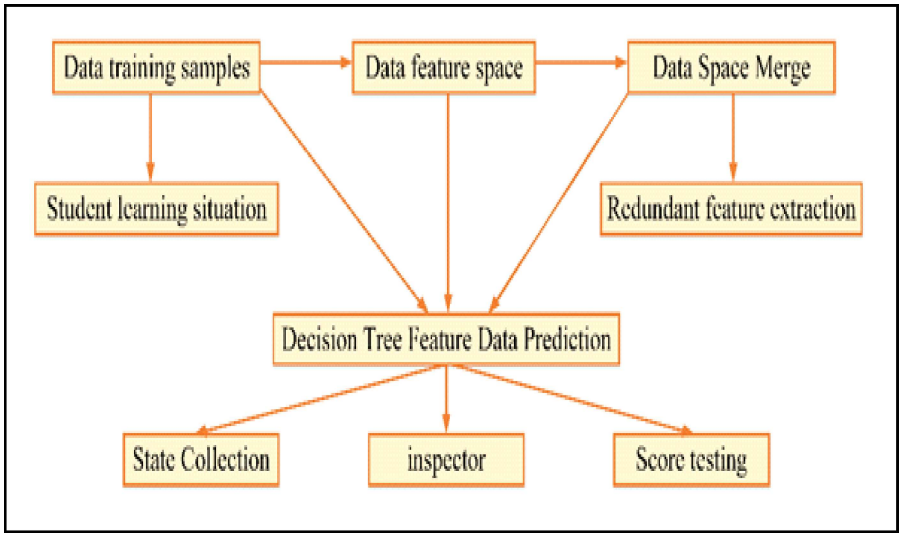

Figure 1. Decision Tree English Speaking Proficiency Prediction Model Structure

Figure 1 shows that dynamic variables of students' learning progress are incorporated during the training sample process. Effective features are selected based on the feature space and difference calculation method. Finally, when evaluating the accuracy of the test results, the changing state of the previous data sequence must be considered.

3.2 English Speaking Proficiency Prediction Based on Decision Tree Algorithm Optimization

In the English subject learning of vocational college students, the importance of English speaking is often overlooked. Both teachers and students lack the construction of English-speaking contexts, and although English-speaking exercises are interspersed in many teaching activities, the actual learning outcomes are not satisfactory. Teachers cannot correct students' oral language individually in classroom teaching. With the increase in

pronunciation errors, students' English-speaking proficiency will only deteriorate, gradually losing this means of international communication. Therefore, teaching English presents more challenges than teaching other skills. Teachers not only need to adopt effective solutions but also adjust the classroom structure to provide innovative ideas. Strengthening pronunciation in English speaking is the basic operation to improve speaking proficiency. Teachers can adjust students' pronunciation focus through demonstrations. Guiding students to read words and texts fluently is a simple demonstration process. However, simple reading and recitation cannot meet the requirements of improving fluency. Pre-recording standard English-speaking audio can be used to correct students' oral issues. Additionally, emphasising pronunciation can help students understand the key points of correct pronunciation. We analyzed the effectiveness of these two methods in improving English-speaking proficiency in practical situations, using decision tree optimization to study students' actual English-speaking proficiency and record changes in their English-speaking performance.

We sampled and statistically analyzed the different feature points of the two learning methods and conducted decision tree optimization training in the reconstructed high-dimensional feature space. Useless information was removed, and mainstream data reflecting students' learning progress was extracted as feature variables. Assuming that the data exhibits nonlinear changes, the decision tree model primarily expresses the evolution of the data set in tuples:

$$(R_{m-1}, E_t, E_d, D_t) \tag{3}$$

Among them, R represents the valid nodes in the original data space. By setting the accompanying duration for other interactive data and using the decision tree method for feature distribution training, the actual result coefficients of the two data sets are calculated. Comparing the result coefficients, it is found that strengthening oral skills has a greater impact on developing English-speaking proficiency in vocational college students. Therefore, the effectiveness of the decision tree algorithm in predicting discrete dynamic data is evident, as it can handle the fitting problem between the original data and new data. Using the principles of inductive algorithms to generate decision tree rules accessible to researchers ultimately provides a reliable basis for predicting situations in data analysis.

4. English Speaking Proficiency Prediction System Based on Decision Tree Algorithm Optimization

The emphasis on English speaking in English teaching is relatively lacking, and improving the optimization of English-speaking teaching in vocational colleges and among teachers is a key way to enhance students' overall English proficiency. First, starting from students' learning interests, students' interest in learning English speaking is a prerequisite for their active engagement in educational activities and an essential means to improve learning efficiency. Teachers and schools should help students build confidence in English speaking, clarify their learning goals, and conduct teaching activities more effectively. At the same time, creating a relaxed and pleasant English communication context reduces students' anxiety about the English learning environment. According to the characteristics of teaching and English subjects, students often reduce their practice of English speaking due to a lack of confidence and other psychological factors. Increasing students' oral practice activities in the learning environment can change the inhibition effect of the original language structure. Fully tapping into students' initiative, guiding them to practice their oral skills actively, express their ideas and behaviors, and expand the frequency of vocabulary and grammar usage.

The decision tree algorithm is a primary technique suitable for data classification and prediction. This inductive

method is highly effective for dealing with disorderly and irregular instances. Using a top-down recursive approach, comparing nodes with the same attributes within the decision tree is possible. In the above research, we used decision tree predictions to evaluate the actual effects of the two ways to improve students' English-speaking proficiency, as shown in Figure 2.

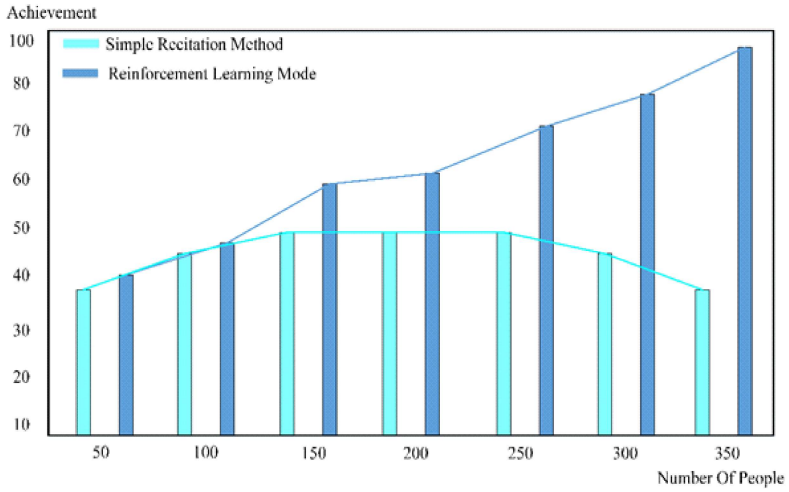

Figure 2. English Speaking Proficiency Variation Chart under Simple Repetition and Reinforcement Learning

Figure 2 shows that the actual performance of English-speaking proficiency in the simple repetition group is inferior to that of the reinforcement learning group. Although the decision tree algorithm provides relatively accurate predictions, an optimized decision tree model is still needed to increase the speed of prediction feedback. We utilized the association algorithm to adjust the prediction structure, reordering and combining the calculation rules of the prediction model, and explored the feedback efficiency of the two algorithms on the same time series.

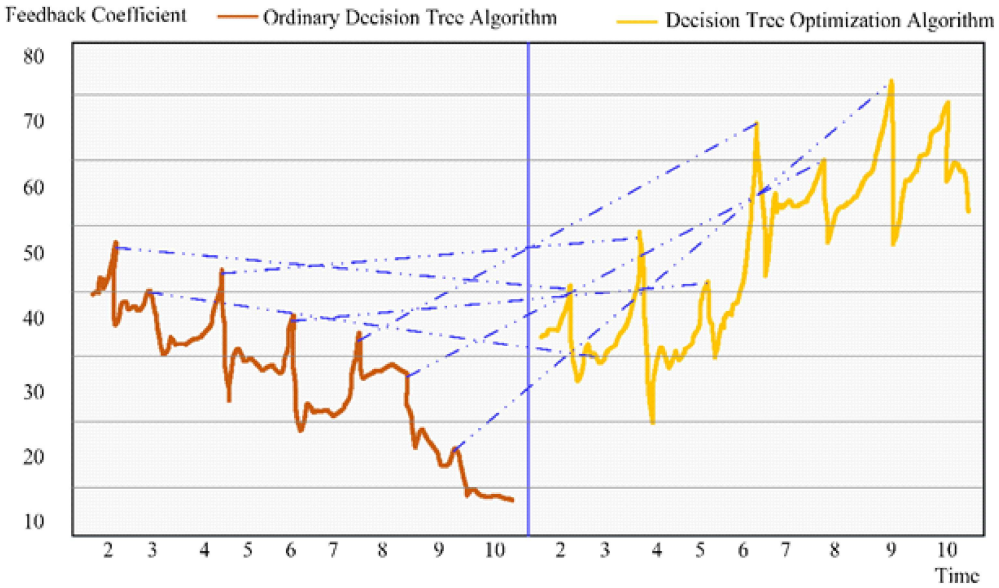

Figure 3. Comparison Chart of Feedback Efficiency between Two Decision Tree Algorithms

Figure 3 shows that the conventional decision tree algorithm is significantly slower in predicting students' English-speaking proficiency compared to the optimized decision tree algorithm. At the same moment, the decision tree algorithm optimized with association rules provides more feedback information. In addition, we can also use the classification of continuous attributes to verify the algorithm's effectiveness. By adopting standardized methods, we judge the algorithm's reliability to provide a basis for subsequent assessment of students' actual English-speaking situations. In practical application, we incorporate students' speaking practices into the performance evaluation system. Based on the students' vocal characteristics, we create voice waveform charts to compare and evaluate the quality and fluency of their spoken English, thus providing fundamental assessments and judgments.

5. Conclusion

With the development of economic globalization, research on English teaching levels in the education field has gained increasing attention. More and more English teaching focuses on practical training rather than theoretical instruction, with English speaking proficiency being the most crucial aspect of applied skills and receiving extensive attention from educational researchers. The English-speaking proficiency of students in vocational colleges is an essential basis for evaluating their English skills and levels. This study employed the decision tree algorithm to assess students' English-speaking proficiency, supported by predictive analysis of actual speaking practices. Firstly, we analyzed the current development of students' English-speaking proficiency in the big data environment and then selected the decision tree algorithm to construct the predictive model for students' speaking proficiency. By optimizing the decision tree algorithm with association rules, we improved its efficiency in prediction and calculation. To verify the reliability of the decision tree algorithm in the predictive model, we incorporated multiple assessment perspectives to measure the variations in students' English-speaking scores. The research results demonstrate that the optimized decision tree algorithm accurately predicts actual English-speaking proficiency among students in vocational colleges and provides a certain basis for reliability.

Acknowledgement

The 2023 Henan Province Vocational Education Demonstration Professional Point Construction Project supported the study.

References

[1] Wu, C. (2023). The Impact of Positive Psychology-based Classroom on English Speaking Teaching and Learning of EFL College Students. *Journal of English Language Teaching and Applied Linguistics, 5*(3), 45-54.

[2] Liao, H., Li, L. (2023). Facilitating EFL learners' intercultural competence through culturally responsive teaching in oral English classrooms. *System, 103070*.

[3] Rahimi, M., Ong, K. K. W. (2023). Exploring expert teachers' cognitions and practices of teaching English speaking and their students' experiences and engagement. *System, 103064*.

[4] Kitjaroonchai, N. (2012). Motivation toward English language learning of students in secondary and high schools in education service area office 4, Saraburi Province, Thailand. *International Journal of Language*

and Linguistics, 1(1), 22-33.

[5] Zhou, H., Yu, Q. (2023). The Study of Opening and Closing of Conversations in Oral English Textbooks. *Journal of English Language Teaching and Applied Linguistics, 5*(2), 183-190.

[6] Roengrit, P., Wathawatthana, P., Hongsa, N. (2023). Improving the English Speaking Ability of Sixth Grade Thai Students Using the Role-play Technique. *English Language Teaching, 16*(6), 176-176.

[7] Dai, X., He, B., Lu, X., et al. (2023). Research on the Application of Formulaic Language in College Students' Spoken English Communication. *Lecture Notes on Language and Literature, 6*(3), 20-26.

[8] Hu, M., Guo, L., Liu, J., et al. (2023). Classification and Recognition of Building Appearance Based on Optimized Gradient-Boosted Decision Tree Algorithm. *Sensors, 23*(11), 5353.

[9] Barros, B., Conde, B., Cabaleiro, M., et al. (2023). Design and testing of a decision tree algorithm for early failure detection in steel truss bridges. *Engineering Structures, 289*, 116243.

[10] Mehrpour, O., Hoyte, C., Nakhaee, S., et al. (2023). Using a decision tree algorithm to distinguish between repeated supra-therapeutic and acute acetaminophen exposures. *BMC Medical Informatics and Decision Making, 23*(1), 102.