Design of Information Security Evaluation Model in English Teaching under Embedded Network

Jun Liu[1,2] ,Xiao-ai Niu[1*]
[1]Shaanxi Technical College of Finance and Economics; Xianyang, Shaanxi, China
[2]Key Laboratory of Optoelectronics, Xizang Minzu University; Xianyang, Shaanxi, China
**Corresponding Author*

Abstract:

This article thoroughly and comprehensively explores the innovative application of embedded neural network technology in the multimedia teaching mode of universities for automatic scoring and feedback design in English teaching. Based on this, a scoring feedback generation system that can be seamlessly integrated into the actual teaching process is carefully designed. A major breakthrough of this system is that it creatively integrates the mature technology of the validated English composition automatic scoring system with the most advanced speech recognition engine, jointly constructing a multifunctional intelligent scoring model with embedded neural network as the core. This model is specifically designed for open-ended oral problems, aiming to provide an efficient and accurate automatic scoring solution.

Keywords: college English; multimedia teaching; automated grading; feedback generation system; Embedded Neural Network; Information Security

1 INTRODUCTION

The traditional manual marking method, although it has long been the main means of evaluation, its inherent limitations are becoming increasingly prominent. It is not only time-consuming and inefficient, but also difficult to ensure fairness in the marking process, easily influenced by personal subjective factors, which leads to widespread questioning of the fairness of scoring. In contrast, the introduction of automatic grading technology provides an effective solution to this problem [1]. Automatic grading not only significantly improves grading efficiency and shortens the grading cycle, but more importantly, it can ensure the consistency and fairness of subjective question grading through preset algorithms and standards, meeting the urgent public demand for fairness and transparency in the grading process. Although there are various scoring systems in the current market, most of them are still limited to automatic scoring of objective questions. It is particularly difficult to achieve automatic grading for question types such as English Chinese translation, which have strong subjectivity, broad knowledge, and open answers. This is mainly because such questions not only require evaluators to have a strong language foundation, but also to have flexible thinking and judgment [2]. However, current English teaching often focuses too much on imparting language knowledge, neglecting the subject status of students in the reading process, resulting in difficulties in forming a systematic conceptual framework for students' reading experience and a lack of extensive transfer and application abilities. It is worth noting that structured questions without standard answers, such as open-ended discussions and case studies, are of great significance for cultivating students' creativity and problem-solving abilities [3]. However, the proportion of such issues in high-risk testing is relatively small, and their educational value has not been fully realized. In order to cultivate students' deep thinking ability in reading teaching, teachers need to actively guide and encourage students to engage in reflective evaluation, not only evaluating their own thinking process, but also learning to evaluate others' perspectives, thereby promoting the deepening and expansion of thinking.

Embedded neural networks, as an important technology in the field of deep learning, have shown great potential for applications in multiple fields due to their powerful data processing capabilities, efficient feature extraction mechanisms, and adaptive learning characteristics [4]. In the field of education, embedded neural networks can not only accurately capture students' learning behavior characteristics, but also deeply understand the complex structure of learning content, thus providing strong support for personalized teaching. For multimedia teaching of college English, traditional scoring and feedback mechanisms often rely on teachers' subjective judgments or simple quantitative indicators, which are difficult to comprehensively and accurately reflect students' real learning status and language proficiency. The introduction of embedded neural networks provides the possibility for building more scientific, objective, and intelligent scoring and feedback systems. Through deep mining and analysis of massive data in students' learning processes, embedded neural networks can automatically identify students' learning patterns, difficulties, and strengths, and generate personalized scoring and feedback suggestions based on this, thereby helping students better understand themselves and improve their abilities.

Therefore, this study aims to explore the application of embedded neural networks in the grading and feedback design of multimedia teaching in college English. By constructing an intelligent grading and feedback system based on embedded neural networks, comprehensive monitoring and accurate evaluation of students' learning process can be achieved, providing strong support for the improvement of the quality of college English teaching. At the same time, this study also hopes to provide valuable experience and inspiration for teaching reform in other disciplines through practical exploration.

2 RELATED WORKS

Kim et al. analyzed the changes in four aspects of vocational education in the era of artificial intelligence: economy, policy, ecology, and wisdom, and further pointed out the opportunities and dilemmas faced by vocational education [5-8]. Prasetya discusses the automatic criticism of spoken language based on the composition automatic criticism system [9]. It is feasible to combine speech recognition technology and composition criticism system for spoken language criticism, but the criticism method for spoken language proposed in the paper still has limitations, its scoring dimensions have great limitations and it does not realize the evaluation of speech itself. In this paper, Zhang et al. have done some pioneering and important work on the automatic scoring of retelling questions with the help of speech recognition and "word map" based machine scoring feature extraction, and successfully scored retelling questions automatically [10]. However, the retelling questions in Zhang et al.'s work are still different from the oral questions in this paper [11]. In addition, the scoring process also relied on the data of the test takers' reading questions and did not fully realize the separate scoring of the retelling questions.

In the challenge of scoring open-ended oral questions, traditional generic language model segmentation methods are inadequate due to the diversity of answers and the lack of fixed expert models or keyword annotations. However, inspired by the research of Parsazadeh et al., especially their innovative method of representing speech recognition results through word graphs, we can use embedded neural networks to construct a more flexible and powerful scoring system. This article focuses on improving the accuracy of the model in various aspects of scoring, especially when dealing with college English oral assessment [12]. In the preprocessing stage, it is crucial to finely denoise and cut the silent segments in the speech file, which not only improves the accuracy of speech recognition but also serves as the foundation for subsequent analysis. To this end, we will conduct in-depth research and apply existing noise reduction techniques, especially combining traditional noise reduction algorithms with deep learning, to develop advanced noise reduction solutions suitable for automatic speech evaluation [13].

In the speech recognition stage, we will use an open and free speech recognition engine for initial transcription, but we are aware of the limitations of current technology, namely that the recognition rate has not yet reached perfection. Therefore, before grading, we must conduct a deep cleaning of the recognized text to correct grammar and word errors caused by pronunciation or fluency issues, in order to more accurately reflect students' actual expressions. Here, embedded neural networks will play a crucial role in distinguishing true language errors from recognition errors through their powerful pattern recognition capabilities.

3 ANALYSIS OF THE AUTOMATED GRADING FEEDBACK SYSTEM OF ENGLISH MULTIMEDIA TEACHING MODE IN COLLEGES AND UNIVERSITIES

3.1 Automated scoring design for multimedia teaching mode

In the task of comparing time series similarity, the two-time series are usually unequal in length. For example, in speech recognition tasks, the speed and pitch of human speech may vary, resulting in different pronunciation speeds of different phonemes when saying the same word, e.g., sometimes a sound that should be long is pronounced very short, or vice versa [14]. Besides, the two-time series may just not be aligned on the time axis, while the two-time series themselves are identical, and aligning them only requires shifting one of the segments on the time axis. Especially in speech recognition tasks, the distance between two speech sequences cannot be effectively found using the traditional Euclidean distance, because the case of two-time sequences may be very complex.

Students' answers to test papers are all on their answer cards. Nowadays, paper marking is not practical in larger exams, and the online marking method is to scan students' answer cards into answer pictures for marking. To facilitate the marking of test papers by teachers, it is necessary to cut the answer card pictures into certain blocks by topic, and the marking teacher will only see the cutout part of the test paper that the teacher is responsible for correcting. Enter the information about the test to be corrected into the system, including but not limited to the name of the test, the time of the test, the marking rules of the test questions (each question has a maximum mark, whether it can be given 0.5 marks, etc.) and the basic information of the students, etc. After all the previous preparations are completed, the marking instructor grades the blocks of questions that he/she is responsible for marking, and the marks given need to satisfy the marking rules of the test questions.

One of the goals of automatic scoring is to achieve a uniform and fair scoring of each question type. The image obtained from the scanned answer card contains several blocks of student responses, and the automatic scoring method must be different for different blocks of questions. Therefore, it is still necessary to use the test cutter to split the answer card image into question blocks.

The automatic scoring system finally analyzes the text data and gives the score, and the current student response images need to be transformed to get the student response text, so the automatic scoring process needs to add the text transformation step to the traditional marking process. Currently, there are two ways to achieve this: firstly, the electronic image of the answer card is cut by the test questions to get the picture of each question block, and then the picture of each question block is converted into the text data of each question block by the graphic transformation; secondly, the electronic image of the answer card is firstly converted into the text data of the answer card, and then the text data of the answer card is cut by the test questions to get the text data of each question block. Considering that image batch cutting is easier than text batch cutting and the chance of error is smaller, the former is used, as shown in Figure 1.

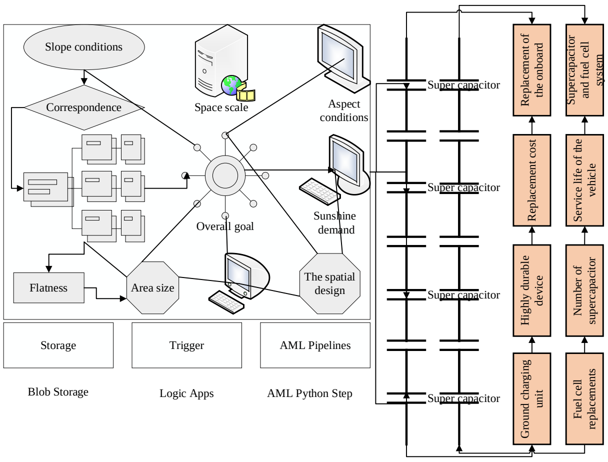

Figure 1 Multimedia automated scoring model

Among them, the variety and complexity of sentence patterns and the quality complexity of words are the key emphases, i.e., the composition writing process cannot be all simple sentences; overly simple sentence patterns and simple vocabulary can make the quality of sentences much worse [15]. Therefore, the variety of sentence types, sentence complexity, and advanced vocabulary can reflect the good or bad quality of sentences. Given that the primary goal of this chapter is to establish a sentence elegance recognition model, the quality of sentence form should be considered along with the rhetorical devices, such as metaphorical references, prose, contrast, etc., to evaluate the elegance of sentences comprehensively.

Using the NLTK toolkit, the text is divided into words, lexical annotation, root reduction, and case unification, and then the average word length and the total number of words can be counted, the number of advanced words in the sentence can be counted by matching the advanced vocabulary, the lexical annotation can be used to obtain the lexicality of each word, and then the number of conjunctions in the sentence can be counted.

The syntactic analysis of the sentences was performed using the Stanford Parser toolkit, and the syntactic tree was constructed, and the tree structure is shown in Figure 2. The syntactic analysis can count the number of phrases in the sentence, including noun phrases (NP), verb phrases (VP), prepositional phrases (PP), adverbial phrases (ADVP), etc. The number of subordinate clauses (IP) of the sentence can also be counted, as shown in Figure 2.

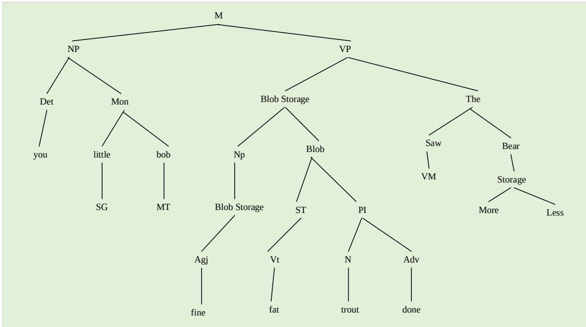

Figure 2 Example diagram of the syntactic tree structure

The purpose of this section is to use embedded networks to extract beautiful features from sentences, which requires us to achieve a transformation from word granularity features to sentence granularity features. In the input layer of embedded networks, we use Word Embedding technology to convert each word into a vector in a high-dimensional space. These vectors together form a two-dimensional vector matrix M × K that can represent a sentence, where M is the number of words in the sentence and K is the dimension of each word vector.

$$C_j = f(w * x_{i,i+j-1} - b_j) \quad (1)$$

$$f(x) = \min(0, x^2) \quad (2)$$

The convolution kernel is applied to the input data to obtain a feature map C, where:

$$C = \left[C_1, C_2, ..., C_m\right] \quad (3)$$

$$sim(v,u) = \frac{vec(u) \otimes vec(v)}{|vec(u)| \cdot |vec(v)|} \quad (4)$$

The word-level similarity between composition texts and composition topics is discussed to characterize the relevance between composition content and composition topics. In this section, we extract sentence-level correlation features to further analyze the degree of relevance of the composition by introducing the concept of semantic dispersion and analyzing how closely the sentences are related to the topic. Semantic dispersion refers to the degree of difference between the semantics of sentences. In the AES problem, we split the composition into several sentences [16]. Then the semantic information difference between the sentences and the composition title and the semantic information difference between the sentences and the composition text are calculated separately. The former identifies how closely each sentence is related to the essay topic, and the latter identifies how closely each sentence is related to the essay text. The semantic dispersion is calculated by combining the two aspects, and the degree of thematic relevance of the sentences in the composition is measured at the sentence level.

$$d_k = \left[d_{k,1}, d_{k,2}, ..., d_{k,n}^2\right] \quad (5)$$

The formula for each dimension d, in this vector, is shown in (6).

$$d_{k,\mathrm{n}} = \frac{\arcsin\left(\frac{vec(u) \otimes vec(v)}{|vec(u)| \cdot |vec(v)|}\right)}{2\pi} \quad (6)$$

Features based on sentence granularity have the greatest influence on the scoring results, followed by word granularity features, and chapter granularity has the least influence. This is because chapter granularity characterizes text similarity by constructing text vectors throughout the extraction, which is coarse in terms of feature extraction, while word granularity characterizes text similarity by the similarity between word vectors, which is more refined and does not grasp the global semantic information, while the sentence granularity-based features well to combine the advantages of the first two feature extraction methods.

When building an automatic scoring model based on embedded networks, we make full use of the textual information reflected by students, the scoring criteria information, the reference answer information after graphic conversion and proofreading, and the corresponding student grade information. These diverse data sources provide a solid foundation for model training. In order to ensure that the scoring model accurately reflects the standards of professional raters, we emphasized the integration and preprocessing of this data during the training phase, in order to store and manage it in a unified format, paving the way for subsequent model training and automatic scoring processes. In the data preprocessing stage, we carefully cleaned, converted, and standardized the student response text, grades, and scoring criteria to ensure the accuracy and consistency of the data. Subsequently, based on manual scoring criteria, we carefully selected features that can characterize the quality of students' answers and constructed a scoring feature set. These features not only cover multiple dimensions of text content, but also fully consider the specific requirements of scoring criteria, providing rich information input for embedded network models. In the model training phase, we chose embedded networks as the core architecture and utilized their powerful feature extraction and representation learning capabilities to explore the complex relationship between response text feature sets and scoring datasets. By continuously optimizing the network structure and parameters, we aim to train an English Chinese translation question scoring model that can accurately predict students' answer scores. In the testing phase, we also extract corresponding features of students' reaction texts and use a trained scoring model to calculate these features, in order to quickly and accurately provide scoring results. The entire scoring process is highly automated, greatly improving the efficiency and accuracy of scoring.

3.2 Design of English feedback generation system in higher education

Requirements 3, 4, 5, 6 and 7 are to maintain and manage information related to grading, including the management of grading models and automatic grading functions; Requirement 8 is to maintain and manage user login information. Requirement 8 is the maintenance and management of user login information [18]. Thus, the English-Chinese automatic marking system can be divided into four modules: one is the login information management module, the other is the test information management module, the third is the student information management module, and the fourth is the marking information management module, and the module design of the English-Chinese automatic marking system is shown in Figure 3.

The login management function is mainly for the identity verification of the super administrator user in this system, the user needs to enter the user's name and password and click the login button, the system will compare the super administrator user information in the database and verify the information entered by the user. If the verification information is correct, the user will enter the system directly, otherwise, the user's name or password will be prompted with an error.

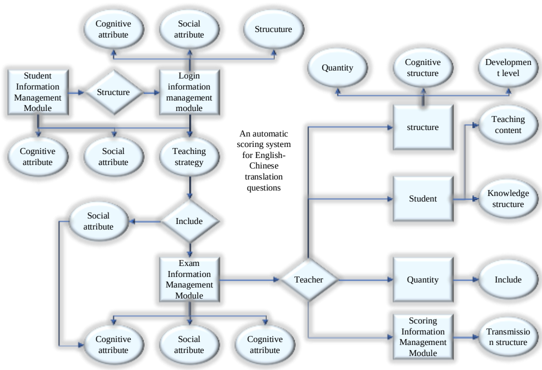

Figure 3 System module design diagram

Once the user has successfully logged in, he/she can manage the password information and replace the old password [19]. Again, you need to enter the password to verify the user's super administrator identity and enter the new password information to be changed. The system will judge whether the two password information entered meets the format size requirements. If the password information meets the requirements, then the password change is successful and the login screen is returned, otherwise, it will prompt that the password information is wrong and needs to be input again for verification.

In large-scale college entrance exams, the markers usually give scores in the form of double or multiple markings, i.e., two or more teachers mark the same response text together. In the model training process, student responses and student scores are one-to-one correspondence, so it is necessary to normalize the information collected from the response text corresponding to the scores [20].

The lexical length characteristics of the translated sentences were selected above to simply examine the internal lexical structure of the translated sentences, without comparing them with the reference answers. The lexical match between the translated text and the reference answer is now considered to be included. Since the 1-4 tuple match rate has already been calculated when examining the text overlap previously, and to keep the translated sentences somewhat flexible, the lexical five-tuple match rate between the two is extracted for quantitative representation. The automatic scoring feature terms of the English-Chinese translation questions analyzed above are organized and summarized, and the results are shown in Figure 4.

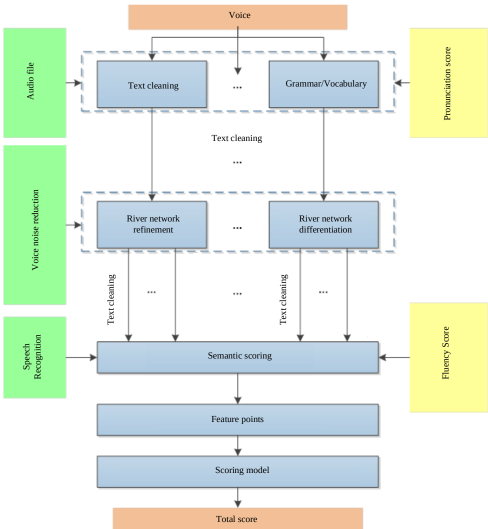

Figure 4 Multi-feature intelligent scoring model

The exam management function is mainly for adding, deleting, and modifying exam information. Users can add an exam information record to the database by entering the relevant information of the exam. Click the Modify button to modify the name and date of the exam, or click the Delete button to delete the exam from the database. The question management function is also divided into adding questions, deleting questions, and modifying questions. Users enter the question number, the corresponding question type (English to Chinese or Chinese to English), the exam to which it belongs (previously added in the test management), the full score value, and the reference answer, and the system will verify the added question information by clicking the OK button. You can also click the Modify button to modify the question number, question type, exam, full mark value, and reference answer for the successfully added question, or click the Delete button to delete the question from the database and prompt the corresponding error message if the question information is wrong.

In the case of double scoring, the average is taken, and multiple scoring is selected according to the corresponding ratio so that the final score information corresponding to the response text may not be an integer and needs to be appropriately traded off according to the full score value of the current question type and the scoring standard information.

In terms of validity, the instant feedback from the review website is effective in increasing students' interest in writing, revising their essays repeatedly, and thus improving their writing skills. However, due to the limitations of the technology of the web-criticism website, the feedback given to students at the level of vocabulary and grammar is more precise because the scoring rules for vocabulary and grammar are relatively clear, in which vocabulary is mainly examined for richness and complexity, and grammar is mainly examined for correctness and richness of grammatical structures [21]. However, the grading website does not yet give students enough feedback on writing content, chapter structure, style rhetoric, content logic, and coherence, all of which are relatively subjective. Therefore, only two of these scores, vocabulary, and grammar, are extracted in this paper.

4 ANALYSIS OF RESULTS

4.1 Automatic analysis performance analysis

In order to comprehensively and scientifically evaluate the multi feature intelligent criticism model based on embedded networks proposed in this paper, we introduce a rigorous quality assessment method. Given that the task of this model in oral criticism is essentially to simulate and approximate the process of manual criticism, the evaluation criteria for its performance are naturally closely linked to the results of manual criticism. Combining the advantages of embedded networks, this model has demonstrated strong potential in various aspects such as feature extraction, representation learning, and rating prediction. By continuously optimizing the network structure, adjusting parameter configurations, and introducing richer training data, we can further improve the performance of the model, making it closer to or even beyond the accuracy and consistency level of manual criticism. Therefore, in the evaluation process, we will also focus on the contribution of embedded networks to the improvement of model performance, in order to provide strong support for future model improvement and optimization.

To suggest the next activity, the system tries to match the current sequence with the A-B-C-B-B-D activity sequence by aligning the two sequences and then shifting the design sequence. After the third shift, an activity matching the existing C-B-B sequence is found, and then the next activity D of the known sequence is added as a candidate. If no match is found, the window size is reduced by removing the first activity of the input queue to increase the likelihood of finding a matching sequence in the existing pattern, as shown in Figure 5.

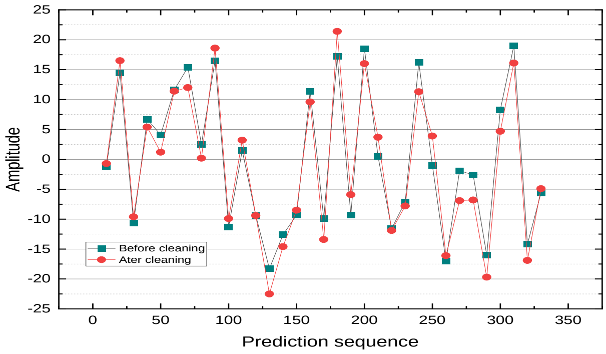

Figure 5 Comparison of scores before and after text cleaning

From the state transfer network, we can see that A and B have been used as the starting point of teaching activities in the teaching case; all activities except A, B, G, and M have been used as the endpoint; and C, D have been used as the starting point and have been designed as the end point of activities. The connecting line between two nodes indicates the transfer relationship between activities in the instructional case, and the numbers on the connecting line indicate the probability of transferring one state to another state (containing itself). Further analysis of the state transfer network leads to the following conclusions: there is no certain sequentially stable relationship within these teaching cases, and there is no fixed transfer pattern between learning activities.

As shown in Figure 6, the scoring model trained with DNN classification achieves 89.3% of the performance of manual scoring, which exceeds the traditional linear regression model, so the scoring model can be practically applied in the examination system. In the index of mean score difference, the DNN model is also superior to linear regression, so the performance of the scoring model can be effectively improved by using the DNN classifier to train the scoring model.

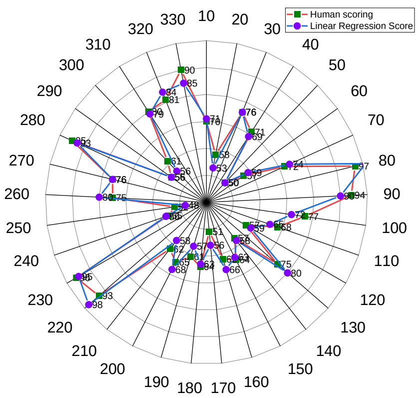

Figure 6 Automatic scoring model scoring results

According to the manual scoring criteria of open-ended speaking questions, we propose a multi-featured intelligent speaking review model. This multi-featured automatic review model is fully compatible with the manual scoring criteria, and the five dimensions of English speaking: grammar, vocabulary semantics, pronunciation, and fluency are scored comprehensively, and the scoring dimensions are more comprehensive and accurate than the open-ended speaking automatic scoring of related studies. The reliability of the system was improved in terms of the scoring dimensions. The DNN classifier was used to train the scoring model: compared to the linear regression model often used in previous studies, the DNN classification model improved the correlation between machine scoring and expert scoring, which led to a new level of accuracy of the automatic open-ended speaking scoring. Also, there is a slight improvement in the correlation between human and machine scoring relative to others' studies (88.5%).

4.2 Analysis of system test results

The first round of the testing process was mainly conducted online, with the help of platforms such as Pinning and StudyTalk to carry out the film and television choreography and production course. The teacher carries out teaching through the online platform, and the main contents include general knowledge, case explanation, task assignment, and other operations. In this round of testing, the author and the instructor jointly completed the instructional design. Since the main practical devices for online students are students' smartphones and tablets, which exclude interference in offline courses, students can be in a teaching environment that they are familiar with and can conduct independent learning, which can improve learning efficiency.

According to the Torrance Creativity Evaluation Scale, the subjects can be considered to have average creativity when the scores are 0-13, good creativity when the scores are 14-17, and very good creativity when the scores are 18-20. According to the test results, it can be found that the overall creativity level of the students improved after the first 12 hours, which can be considered as the improvement of the cognitive ability of the students in the film and television choreography course itself. The increase in students' overall creativity level at the end of the course was higher than the increase in students' creativity level at the end of the first 12 credit hours, indicating that students' cognitive level was significantly improved after using the research-designed instructional design automation tool for learning activity-centered instruction, as shown in Figure 7.

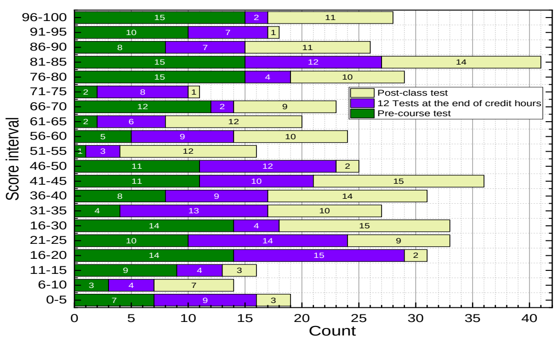

Figure 7 Test results of Torrance Creativity Rating Scale

As can be seen from the figure, the overall scores of students in the third and fourth stage tests showed a large increase compared to the first time, indicating that students' proficiency in the theoretical knowledge has improved significantly, which indicates that the use of the instructional design automation tool helped the teacher to better organize the classroom and students' learning efficiency has improved. At the end of the course, a questionnaire was administered to some of the participating students through the online learning tool. The main questions of the survey included whether they were satisfied with the course content arrangement, the degree of student participation in the classroom, and the degree of application of the technology tools recommended by the instructor.

The experimental dataset used in this project is from a large English test in China and contains a total of 647 English to Chinese questions and the corresponding scoring data, and 912 Chinese to English questions and the corresponding scoring data. The scoring data in this dataset were provided by professional teachers who teach on the front line, which ensures the professionalism and reliability of the data. The full scores of both English to Chinese and Chinese to English questions are 2, and the scores of students' responses are generally between 0.5 and 1.5, with few scores being too low or too high, as shown in Figure 8.

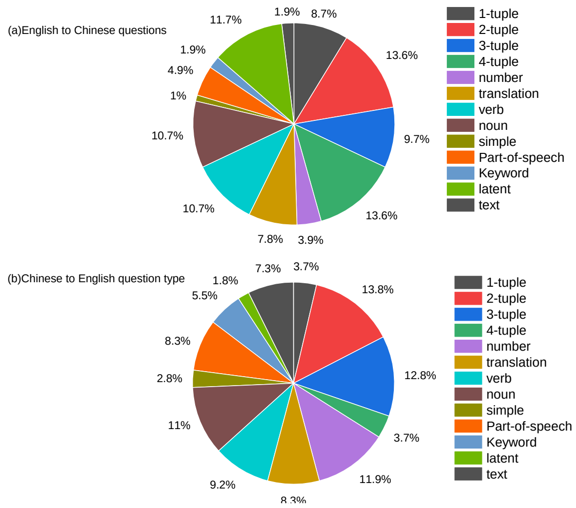

Figure 8 Correlation of scores for each feature term

First, the feature with the highest correlation is the length of the translation in both English to Chinese and Chinese to English, indicating that the completeness feature represented by the length of the translation is not only intuitive to the completion of the translation but also correlates quite well with the actual marking score. The correlations of the scores of the 1-tuple match rate, 2-tuple match rate, 3-tuple match rate, and 4-tuple match rate were all similar in the Chinese-to-English question type, and the correlation of the scores of the 1-tuple match rate was significantly higher than the other three items in the English-to-Chinese question type, indicating that the core vocabulary in Chinese is relatively easier to express the meaning of the text.

The differences in expressions due to the differences in Chinese and English languages, such as the emphasis on structure in English and semantics in Chinese, and the sentence structure features of verb and clause length, are particularly evident in the English to Chinese translation and Chinese to English translation. Finally, there are several semantic-related feature items, all of which show a high correlation, and also indicate the crucial importance of semantics in translation questions. In conclusion, the rating correlations of each feature item of the English-Chinese translation question types selected in this study are all good, but the differences in expressions due to different language types make the influence of each feature item in both English-Chinese and Chinese-English translation question types different, and it is more reasonable to construct rating models for them separately to make the model learn the corresponding features better.

5 CONCLUSION

This article comprehensively and deeply explores the innovative application of embedded neural network technology in the scoring and feedback design of college English multimedia teaching environment, and based on this, successfully designs and implements an efficient and intelligent scoring feedback generation system. This system not only cleverly integrates the mature mechanism of the English composition scoring system that has been tested over time, but also breaks through by incorporating cutting-edge speech recognition engines, opening up a new path for automated scoring of open-ended oral problems. In the key step of feature extraction, the system exhibits extremely high accuracy and comprehensiveness. It can capture and extract multidimensional feature information from the oral records of candidates in a meticulous manner, covering core elements such as the quality of speech, the richness of vocabulary usage, the accuracy of grammar structure, and the coherence of semantic expression. These characteristic information are like delicate portraits, accurately depicting the oral performance of candidates and providing a solid foundation for the scientific and objective grading work in the future.

ACKNOWLEDGMENTS

This work was supported by 2023 Shaanxi Province Teacher Education Reform and Teacher Development Research Project, Digitally Driven Shaanxi Vocational College Teacher Development and Evaluation Research (No.SJS2023YB10); 2023 Shaanxi Vocational Education Association Education and Teaching Reform Research Project, Digital Economy Era Finance and Economics Vocational College Teacher Development and Evaluation Research (No.2023SZX299).

REFERENCES

[1] Yang X. Automatic recommendation system of college English teaching videos based on students' personalized demands[J]. International Journal of Emerging Technologies in Learning (iJET), 2021, 16(21): 42-57.

[2] Prasetya R. English Teaching Based-Strategy LMS Moodle and Google Classroom[J]. English Education: Journal of English Teaching and Research, 2021, 6(1): 32-44.

[3] Wang X, Zhang D, Asthana A, et al. Design of English hierarchical online test system based on machine learning[J]. Journal of Intelligent Systems, 2021, 30(1): 793-807.

[4] Wang T. A Blended Collaborative Teaching Mode in Language Learning Based on Recommendation Algorithm[J]. International Journal of Emerging Technologies in Learning (iJET), 2021, 16(23): 111-126.

[5] Aljawarneh S A. Reviewing and exploring innovative ubiquitous learning tools in higher education[J]. Journal of computing in higher education, 2020, 32(1): 57-73.

[6] Liu S, Wang J. Ice and snow talent training based on construction and analysis of artificial intelligence education informatization teaching model[J]. Journal of Intelligent & Fuzzy Systems, 2021, 40(2): 3421-3431.

[7] Shen L. Data mining artificial intelligence technology for college English test framework and performance analysis system[J]. Journal of Intelligent & Fuzzy Systems, 2021, 40(2): 3489-3499.

[8] Kim H S, Choi U Y. Learner Perception of an Online L2-Course Summative Exam[J]. Multimedia-Assisted Language Learning, 2020, 23(3): 258-279.

[9] Prasetya R E. The Design of Moodle-Based English Language Learning Environments (Case Study of Indonesian Higher Education)[J]. ELT Worldwide: Journal of English Language Teaching, 2021, 8(2): 222-239.

[10] Zhang Y, Yi D. A new music teaching mode based on computer automatic matching technology[J]. International Journal of Emerging Technologies in Learning (iJET), 2021, 16(16): 117-130.

[11] Zhang R. A Study of the Cultivation of College Students' Learner Autonomy in the Multimedia Environment[J]. Advances in Educational Technology and Psychology, 2021, 5(7): 1-12.

[12] Parsazadeh N, Cheng P Y, Wu T T, et al. Integrating computational thinking concept into digital storytelling to improve learners' motivation and performance[J]. Journal of Educational Computing Research, 2021, 59(3): 470-495.

[13] Smith G G, Haworth R, Žitnik S. Computer science meets education: Natural language processing for automatic grading of open-ended questions in ebooks[J]. Journal of educational computing research, 2020, 58(7): 1227-1255.

[14] Weller J M, Coomber T, Chen Y, et al. Key dimensions of innovations in workplace-based assessment for postgraduate medical education: a scoping review[J]. British Journal of Anaesthesia, 2021, 127(5): 689-703.

[15] Eltahir M, Alsalhi N R, Al-Qatawneh S, et al. The impact of game-based learning (GBL) on students' motivation, engagement and academic performance on an Arabic language grammar course in higher education[J]. Education and Information Technologies, 2021, 26(3): 3251-3278.

[16] Setiawan H. Developing Interactive Multimedia for Teaching Reading Comprehension on Narrative Texts Based on South Sumatera Local Culture[J]. IDEAS: Journal on English Language Teaching and Learning, Linguistics and Literature, 2021, 9(2): 632-641.

[17] Nami F. Towards more effective app-assisted language learning: The essential content and design features of educational applications[J]. Issues in Language Teaching, 2020, 9(1): 245-278.

[18] Jin M. Achievements analysis of mooc English course based on fuzzy statistics and neural network clustering[J]. Journal of Intelligent & Fuzzy Systems, 2020, 39(4): 5559-5569.

[19] Liang X, Haiping L, Liu J, et al. Reform of English interactive teaching mode based on cloud computing artificial intelligence–a practice analysis[J]. Journal of Intelligent & Fuzzy Systems, 2021, 40(2): 3617-3629.

[20] Salih A A, Omar L I. Season of Migration to Remote Language Learning Platforms: Voices from EFL University Learners[J]. International Journal of Higher Education, 2021, 10(2): 62-73.

[21] Chen X. Multimedia teaching system based on art interaction technology[J]. Computer Science and Information Systems, 2022 (00): 26-26.