Integrated Natural Language Processing Algorithms for Automatic Feedback in English Writing: Improving Student Learning Outcomes Through Technology

Yong Liu[1], Xiaoliang Shi[1]
[1]School of Languages and Cultural Communication, Xi'an Mingde Institute of Technology, Xi'an 710124, Shaanxi, China
Corresponding author: Xiaoliang Shi (e-mail: shixiaoliang2000@126.com)

ABSTRACT English writing is a core component of language learning and an important indicator of students' comprehensive language application ability. Traditional writing instruction often faces heavy teacher correction workload, delayed feedback, inconsistent evaluation standards, and insufficient personalized guidance. This study integrates mainstream natural language processing algorithms, including text segmentation, part-of-speech tagging, grammatical error correction, semantic analysis, discourse-structure evaluation, and similarity-based scoring, to construct an automatic feedback model for English writing. A controlled experiment with non-English major college students is designed, with an experimental group using the automatic feedback system and a control group receiving traditional correction. Statistical analysis is conducted on grammar-error recognition, sentence optimization, discourse organization, lexical-collocation correction, writing-score changes, and learning autonomy. The results show that the integrated NLP feedback system can provide real-time and refined revision suggestions, shorten feedback cycles, standardize writing logic, and improve students' writing proficiency and autonomous learning ability. In intelligent education environments, wireless learning terminals, cloudedge data transmission, and educational signal-processing infrastructure further support scalable deployment. The study provides a feasible path for technology-enhanced English writing instruction.

INDEX TERMS natural language processing, automatic feedback, english writing, wireless learning terminals, educational signal processing

I. INTRODUCTION

A. RESEARCH BACKGROUND

WRITING is an important part of language learning, and effective feedback on students' writing is crucial for improving their writing ability [1,2]. In Chinese middle schools, teacher evaluation is the main source of writing feedback for students [3]. Although teacher feedback can accurately indicate areas for revision, teachers' heavy workload and time constraints lead to monotonous feedback content, inconsistent standards, and a lack of diversity and personalization [4]. The traditional education model faces challenges such as unbalanced resource allocation, low teaching efficiency, and difficulty in meeting personalized needs [5].

In recent years, the rapid development of artificial intelligence has profoundly changed all fields of society, including education [6]. With the continuous advancement of AI technologies represented by natural language processing (NLP), artificial intelligence has gradually become an important force driving educational innovation and development [7]. Generative AI feedback has emerged as a new way to assist teacher feedback and improve students' writing ability [8]. Using natural language processing ability, dialogue context understanding ability, and sequential task execution ability, it can easily and efficiently generate comprehensive personalized writing feedback, thereby assisting teachers in providing effective feedback and enhancing students' writing proficiency [9].

Writing feedback refers to the information input that writers receive from readers. Its function is to provide writers with information for revising compositions [10]. It helps students identify problems in expression and revision, gain enlightenment from feedback providers, and continuously improve in idea expression, structure organization, and grammar use [11]. Studies have found that feedback can interact with learners' internal factors to promote their language development [12]. Therefore, writing feedback has positive significance for improving English writing ability.

Compared with traditional artificial intelligence, generative AI has unique technical characteristics. ChatGPT, a generative AI chatbot launched in 2022, uses a deep learning language model to generate human-like text. It works by

analyzing large text datasets and learning patterns and relationships between words and phrases. It can generate natural and diverse responses in dialogue according to context and output human-like feedback [13]. Meanwhile, its enlightening content generation ability, dialogue context understanding ability, sequential task execution ability, and programming language analysis ability can provide various forms of assistance for teachers' teaching. It can also complete multiple tasks of educational evaluation in a targeted manner, and provide targeted feedback by analyzing vocabulary, grammar, sentence structure, viewpoint expression, and event description in the dialogue process [9]. In addition, studies show that compared with direct teacher feedback, human-computer mixed feedback can significantly improve learners' engagement with feedback and thus improve English writing proficiency [14]. Generative AI shows significant advantages in supporting educational feedback.

Therefore, integrating NLP algorithms to realize automatic feedback for English writing can assist teachers in teaching, improve learners' feedback engagement, and make up for the current problems of single teacher feedback forms and lack of personalized feedback. While the cognitive and vocabulary profiles validated in this study specifically address higher education contexts, this framework establishes a baseline that can be incrementally adapted for broader educational levels through age-appropriate prompt engineering and custom corpus filtering.

B. RESEARCH SIGNIFICANCE

(1) Integrate multiple mainstream NLP algorithms to construct an automatic feedback framework suitable for English writing scenarios.

(2) Sort out the application logic of NLP algorithms in text grammar detection, semantic understanding, and discourse structure analysis.

(3) Enrich the theoretical research system of NLP technology in the field of language education.

(4) Quantitatively analyze the impact of intelligent feedback on students' English writing ability and improve the theoretical model of intelligent English writing teaching.

C. RESEARCH STATUS AT HOME AND ABROAD

To contextualize the framework designed to address the aforementioned theoretical and practical significance, it is essential to map the current global landscape of automated corrective feedback. Correcting and giving feedback on students' compositions is a common method for teachers to inform students of their performance and guide revision. It is an important means in the writing teaching process and helps improve students' writing ability. Feedback includes form-focused feedback and content-focused feedback. The former provides feedback on problems in grammar, vocabulary, and writing patterns, while the latter addresses inappropriateness in ideas and logic of article content expression.

At present, domestic and foreign researchers mainly focus on two aspects of English writing feedback: first, the source of feedback. Researchers mainly divide feedback into teacher feedback, peer feedback, and machine feedback based on feedback sources. In terms of teacher feedback, Kepner found that teacher feedback on written texts leads to extensive revisions by students and has a positive impact on English writing. Students who carefully consider and understand teacher feedback show significantly improved English writing quality [15]. In peer assessment, Partridge [16] found that students generally accept peer assessment. Students can learn from each other by revising peers' compositions, improve their own language expression, and enhance their writing confidence to a certain extent.

In addition, domestic and foreign researchers have conducted horizontal comparative and integrated studies on different sources of feedback, making important contributions to exploring new models and ideas for writing feedback. For example, Paulus found that most students benefit from teachers' revisions at the meaning level of their compositions by comparing peer assessment and teacher feedback revisions [17]. Connor & Asenavage found that only 8 college students in the experiment adopted a few suggestions from peer assessment for their own composition revisions, showing low recognition of peer assessment suggestions [18]. Attali applied Criterion in American writing classrooms and found that single scores and total scores of style, grammar, language use, and conventions were significantly higher than the first draft. However, 70% of the students participating in the study were dissatisfied with the feedback from the automatic scoring system [19].

Natural language processing (NLP) is an important research direction in the field of artificial intelligence, with generative AI representing an advanced, contemporary subset within this broader domain. It integrates computer science, linguistics, mathematics, psychology, and other disciplines. Its core goal is to enable computers to understand, generate, and process human language. The framework proposed in this study utilizes a hybrid paradigm: it relies on traditional symbolic and discriminative computational methods for surface-level rule matching and statistical parsing, while leveraging generative large language models exclusively for complex semantic synthesis and contextual feedback generation. In reading education, Kohnke used pre-trained large models to generate high-quality reading texts for reading and writing teaching, significantly improving teachers' teaching efficiency and promoting learners' reading and writing [20]. Su also applied pre-trained large language models to writing teaching but focused on the writing process, using pre-trained large models as automatic evaluators to provide timely and effective feedback for learners' personalized writing learning and solving the defect of lack of effective evaluation and timely feedback from teachers in the past [21].

D. RESEARCH CONTENT AND METHODS

1) Research Content

(1) Sort out the principles of core NLP algorithms and select segmentation algorithms, grammatical error correction algorithms, semantic similarity algorithms, and discourse structure analysis algorithms suitable for English writing feedback.

(2) Integrate various algorithms to build an overall automatic feedback framework for English writing, dividing five functional modules: text preprocessing, error recognition, intelligent scoring, feedback generation, and data statistics.

(3) Design an empirical research plan, select experimental samples, divide experimental and control groups, and formulate a writing evaluation index system.

(4) Make four sets of statistical tables with experimental data to compare and analyze the effect of the automatic feedback system in four dimensions: error recognition accuracy, student writing score changes, error type improvement, and learning autonomy improvement.

(5) Summarize the advantages and deficiencies of system application and put forward optimization strategies for the implementation of NLP algorithms in English writing teaching.

2) Research Methods

(1) Literature research method: Sort out Chinese and foreign literature related to NLP algorithms, English writing teaching, and intelligent feedback systems to lay a theoretical foundation.

(2) Model construction method: Integrate multiple NLP algorithms to build an automatic feedback model for English writing and clarify the operation logic of each module.

(3) Controlled experiment method: Divide students of the same grade and level into experimental and control groups. The experimental group uses the automatic feedback system, and the control group uses traditional manual correction. A one-semester follow-up experiment is conducted.

(4) Data statistics method: Collect data such as students' writing scores, error types, and autonomous revision times before and after the experiment, make standard tables, and conduct quantitative comparative analysis.

II. CORE NLP ALGORITHMS FOR ENGLISH WRITING FEEDBACK

A. TEXT SEGMENTATION AND PART-OF-SPEECH TAGGING ALGORITHMS

English text takes independent words as the basic expression unit and also includes special forms such as abbreviations, fixed collocations, and hyphenated words.

- Segmentation algorithms can automatically split valid words in students' compositions and remove invalid symbols and redundant characters.
- Part-of-speech tagging algorithms can mark corresponding attributes for each word, such as nouns, verbs, adjectives, prepositions, conjunctions, and other common parts of speech.

These two algorithms are the pre-basis for system error recognition. Only by accurately distinguishing words and parts of speech can the system correctly judge whether the grammatical collocation of sentences is reasonable.

B. GRAMMAR ERROR CORRECTION ALGORITHMS COMBINING RULES AND STATISTICS

Grammar error correction algorithms mainly include rule-based algorithms and statistical algorithms.

- Rule-based algorithms rely on a mature English grammar knowledge base and preset judgment templates for tense errors, subject-verb disagreement, number errors, prepositional collocation errors, etc. They can match and identify non-standard expressions in texts.
- Statistical algorithms rely on a large amount of standard English corpus data and judge whether sentence patterns are standard and natural through probability models.

This study combines these two types of algorithms. The combined algorithm can identify basic grammatical errors in students' compositions and detect hidden expression problems such as L1 interference errors (first language interference) and messy sentence patterns.

C. PRE-TRAINED LANGUAGE MODEL SEMANTIC ANALYSIS ALGORITHMS

Pre-trained language models have excellent contextual semantic understanding ability. The algorithm has the following characteristics:

- It can judge the logical relationship between sentences and check the coherence between paragraphs.
- It can judge whether the composition's argument is clear and verify whether the evidence fits the theme.
- It can evaluate whether paragraph transitions are natural and complete in-depth evaluation of the whole article.
- It can provide students with optimization suggestions for discourse logical structure.

D. TEXT SIMILARITY AND SCORING ALGORITHMS

Text similarity algorithms can compare the content differences between students' compositions and standard or excellent model essays. The algorithm has the following characteristics:

- It can calculate the word repetition rate and sentence pattern similarity and judge the fit of the article theme.
- It integrates multiple evaluation dimensions, including lexical richness, grammatical accuracy, discourse integrity, and logical rigor.
- It can build a quantitative scoring model to realize automatic and objective scoring of English compositions.

E. CORE EVALUATION DIMENSIONS OF ENGLISH WRITING

Combined with English teaching standards in primary, secondary schools and universities in China, this paper sorts out five core evaluation dimensions of English writing.

(1) Lexical application: word spelling, word collocation, and lexical richness.

(2) Grammatical application: tense and voice, subject-verb agreement, clause usage specifications, etc.

(3) Sentence structure: collocation of long and short sentences, sentence pattern diversity, and rationality of sentence structure.

(4) Discourse structure: article conception, paragraph structure, logical connection, and effect of beginning and end echoing.

(5) Overall expression: proportion of interlanguage errors induced by L1 thinking patterns, sentence fluency, and content integrity.

F. OPERATION LOGIC OF AUTOMATIC FEEDBACK SYSTEM FOR ENGLISH WRITING

The automatic feedback system supported by NLP algorithms includes five operation steps.

(1) Text input and preprocessing: Students actively input or upload English composition texts. The system automatically completes text segmentation, part-of-speech tagging, and redundant information filtering.

(2) Multi-dimensional error detection: The system relies on various NLP algorithms to check all writing problems from four levels: vocabulary, grammar, sentence pattern, and discourse.

(3) Intelligent feedback generation: The system accurately marks error locations, clearly explains error causes, and provides revision examples and full-text optimization plans.

(4) Automatic quantitative scoring: The system outputs the total composition score and single scores of each dimension according to teaching evaluation standards.

(5) Data statistical analysis: The system automatically records each student's high-frequency error types and generates a class learning status report.

III. CONSTRUCTION OF AUTOMATIC FEEDBACK MODEL FOR ENGLISH WRITING INTEGRATING NLP ALGORITHMS

A. MODEL CONSTRUCTION PRINCIPLES

(1) Practicality principle: The model fits the actual English writing situation of Chinese students, focusing on identifying common problems such as L1 interference syntactic anomalies, basic grammatical errors, and monotonous simple sentences. Feedback suggestions meet classroom teaching requirements.

(2) Integration principle: Integrate multiple NLP algorithms such as segmentation, grammar correction, semantic analysis, and intelligent scoring to realize full-dimensional feedback from vocabulary to discourse.

(3) Simplicity principle: The feedback output by the system is easy to understand, avoiding jargon. Students can directly understand error causes and complete revisions.

(4) Scalability principle: The model reserves interfaces for algorithm optimization and question bank expansion. Model essays can be added, grammar rules supplemented, and deep learning models upgraded later.

B. OVERALL FRAMEWORK DESIGN OF THE MODEL

The automatic feedback model for English writing built in this paper includes five core functional modules: text preprocessing module, multi-algorithm error recognition module, intelligent feedback generation module, automatic comprehensive scoring module, and learning status data statistics module. The five modules are connected and cooperate in sequence. Relying on the integrated NLP algorithms, the model can realize the automatic operation of the whole writing correction process.

C. DETAILED DESIGN OF EACH FUNCTIONAL MODULE

1) Text Preprocessing Module

This module completes basic processing relying on segmentation and part-of-speech tagging algorithms.

(1) The module receives English composition texts submitted by students and first filters invalid information such as punctuation errors, redundant spaces, and garbled characters.

(2) It completes accurate segmentation, splits fixed collocations and professional words, and then tags all words with parts of speech.

(3) The module unifies text formats to provide standardized input data for subsequent error recognition algorithms and ensure analysis accuracy.

2) Multi-Algorithm Error Recognition Module

This module is the core of the model and integrates three types of NLP algorithms to work collaboratively.

(1) The basic grammar layer relies on rule and statistical correction algorithms to detect basic problems such as spelling errors, tense errors, number errors, prepositional collocation errors, and subject-verb disagreement.

(2) The sentence expression layer relies on semantic algorithms to identify problems such as repeated sentence patterns, misuse of clauses, and lengthy and messy sentences.

(3) The discourse logic layer relies on pre-trained language models to analyze deep problems such as paragraph connection, theme fit, and argument logical loopholes.

3) Intelligent Feedback Generation Module

The module generates hierarchical and personalized feedback according to error recognition results.

(1) For spelling and grammatical errors, the system marks error locations and provides correct expressions and concise rule explanations.

(2) For sentence pattern problems, the system provides synonymous revision examples to guide students to replace simple sentences.

(3) For discourse logic problems, the system points out paragraph deficiencies and provides paragraph adjustment ideas and transition sentence examples.

4) Automatic Comprehensive Scoring Module

The module establishes a multi-dimensional scoring index system with a full score of 100: vocabulary (20 points), grammar (30 points), sentence pattern (20 points), discourse (20 points), and expression fluency (10 points).

The higher proportional weight allocated to grammatical accuracy (30%) is strictly derived from the National College English Test (CET) syllabus and localized diagnostic writing rubrics for non-English major undergrad uate freshmen. At this foundational stage of higher education language acquisition, suppressing fossilized structural syntactic errors is instructionally prioritized, thus justifying the higher threshold compared to discourse and stylistic dimensions.

Relying on text similarity algorithms and weighted scoring models, it automatically calculates scores for each dimension and the total score. The system also generates scoring instructions, marks deduction reasons, and helps students clearly understand their writing weaknesses.

5) Learning Status Data Statistics Module

The module automatically records each student's composition submission times, error type distribution, and score change trend. It summarizes class data, counts high-frequency grammatical errors, high-frequency word misuse types, and average writing score ranges. Teachers can accurately grasp the class's overall writing weaknesses through statistical reports generated by the module and reasonably arrange subsequent classroom writing teaching.

D. INTEGRATED APPLICATION PROCESS OF NLP ALGORITHMS

The technical execution across the system architecture operates via a centralized computational pipeline. Rather than executing isolated tasks, the system coordinates the five structural modules synchronously: surface tokenization and syntactic filtration feed directly into a concurrent multi-layered analytical stack where rule-based heuristics and deep-learning semantic language models cross-reference text properties. This unified processing pipeline ensures that immediate quantitative scoring metrics and categorical error feedback profiles are generated in a single end-to-end computational cycle, minimizing systemic lag and maximizing diagnostic alignment.

IV. EMPIRICAL RESEARCH DESIGN AND DATA ANALYSIS

A. EXPERIMENTAL DESIGN

1) Experimental Subjects

This experiment selects two parallel classes of freshmen non-English majors from a university as research subjects, with a total of 96 students: 48 in the experimental class and 48 in the control class. Before the experiment, a unified English writing test confirms that the two classes have no significant difference in average writing level, with the same vocabulary foundation, grammar mastery, and writing ability, meeting the requirements of a controlled experiment. The experiment cycle is one semester, 16 weeks.

2) Experimental Variables

- Independent variable: teaching feedback method. The experimental class uses the NLP algorithm automatic feedback system for composition correction and guidance; the control class uses traditional manual correction by English teachers.
- Dependent variables: students' English writing scores, number of grammatical errors, sentence pattern optimization degree, and autonomous learning ability. It is critical to note that the pedagogical efficacy observed inherently compounds two interconnected dimensions: the structural quality of the algorithmic feedback and the psychological impact of its real-time delivery (immediacy). Consequently, the system is evaluated as a holistic technological intervention—an NLP-driven instantaneous feedback mechanism—acknowledging that the reduction of cognitive latency is a native and inseparable advantage of the algorithmic deployment compared to traditional manual intervals.
- Control variables: The two classes use the same English textbooks, teaching hours, composition training topics, and scoring standards.

3) Experimental Process

(1) Pre-experiment: Organize a unified writing diagnostic test, collect score data, and count initial error types and quantities.

(2) Mid-experiment: Within 16 weeks, both classes complete one English themed composition per week.

- Students in the experimental class submit compositions, and the automatic feedback system provides immediate revision suggestions. Students revise independently according to feedback and resubmit.
- Compositions in the control class are manually corrected by teachers and returned with comments one week later. Students check and revise independently. Teachers provide no extra individual guidance for the experimental class to ensure fairness.

(3) Post-experiment: Organize a unified writing final test at the end of the semester, collect score data, count the number of writing errors and autonomous revision times of students in both classes, and conduct comparative analysis.

B. EVALUATION INDICATORS AND DATA STATISTICS

This paper sets four core evaluation indicators: comprehensive writing score, average number of grammatical errors per person, discourse logic compliance rate, and autonomous writing revision frequency. Based on real data collected before and after the experiment, four standard tables are made: pre-experiment basic score comparison table, pre- and post-experiment writing score change comparison table, average error type quantity comparison table, and student autonomous learning behavior statistics table.

C. EXPERIMENTAL DATA AND TABLE ANALYSIS

The pre-experiment comparison of basic scores between the two classes is shown in Table 1.

It can be seen from Table 1 that before the experiment, the average score of the experimental class is 62.35, and that of the control class is 62.18. The highest score, lowest score, and standard deviation of the two classes are close. The data show that the initial English writing levels of the two classes are basically the same with no obvious gap. The experimental samples are comparable, and the results are valid and reliable. The comparison table of writing score changes before and after the experiment is shown in Table 2.

It can be seen from Table 2 that after one semester of teaching experiment, the average writing score of the experimental class increases by 13.33 points, with an improvement rate of 21.38%. The average score of the control class using traditional manual correction only increases by 6.24 points, with an improvement rate of 10.04%.

The score improvement effect of the experimental class is significantly better than that of the control class. The experimental data fully prove that the automatic feedback system integrating NLP algorithms can effectively improve students' overall English writing learning outcomes.

The comparison table of average error types between the two classes is shown in Table 3.

It can be clearly seen from Table 3 that after the experiment, the average number of four high-frequency errors (word spelling, tense and voice, subject-verb collocation, sentence pattern hybrid) as well as deep discourse logic and argumentation errors in the experimental class is significantly less than that in the control class. The intelligent automatic feedback system can locate various error details and deep semantic inconsistencies in students' compositions in real time and provide corresponding grammar rules, structural optimization maps, and standard revision examples. Students can find their writing problems immediately and complete revision exercises with examples.

The significant reduction in discourse logic and argumentation errors (by 2.03 errors per student) quantitatively demonstrates that the integrated pre-trained language models successfully assist students in rectifying argument alignment and paragraph coherence issues.

The traditional manual correction mode has obvious time delay. Students cannot review wrong questions and consolidate both syntactic and discourse knowledge in time, so they tend to repeat the same writing errors. The statistics table of students' autonomous learning behaviors is shown in Table 4.

The statistical data in Table 4 intuitively compare the learning behaviors of the two groups. The average times of autonomous composition revision, model essay reference, and writing knowledge point sorting in the experimental class are significantly higher than those in the control class.

The intelligent automatic feedback system has the advantage of 24/7 availability. Students are free from time and venue restrictions. After submitting compositions, the system can instantly output complete correction results and optimization suggestions. Students can take the initiative to polish compositions repeatedly according to the system's guidance and actively accumulate writing skills and grammar knowledge. The traditional manual correction mode has a long feedback cycle. Students wait a long time for evaluation results, which reduces their willingness to revise independently and enthusiasm for active learning.

D. COMPREHENSIVE EXPERIMENTAL CONCLUSIONS

(1) The automatic feedback system for English writing integrating NLP algorithms can significantly improve students' comprehensive writing scores, with an improvement range about twice that of the traditional manual correction mode.

(2) The system can accurately identify various grammatical and sentence pattern errors, help students reduce repeated writing errors, and consolidate their language foundation.

(3) Intelligent real-time feedback stimulates students' willingness to learn independently, increases the frequency of composition revision, model essay learning, and knowledge point sorting, and develops good writing learning habits.

(4) The integration of NLP algorithms and English writing teaching has significant application advantages and can effectively improve students' English writing learning outcomes through technical means.

V. EXISTING PROBLEMS AND OPTIMIZATION STRATEGIES OF SYSTEM APPLICATION

A. EXISTING PROBLEMS IN SYSTEM APPLICATION

(1) Limitations in complex semantic recognition: Current NLP algorithms have high accuracy in identifying simple sentence patterns and basic grammatical errors but are insufficient in judging implicit expressions, deep logical ambiguities, and discourse conception of complex argumentative essays.

(2) Insufficient personalized adaptation: The system provides feedback according to universal standards and does not fully consider the differentiated needs of students of different grades and English levels.

(3) Over-reliance on technology weakens teachers' role: Some teachers completely abandon manual comments after

TABLE 1. Pre-Experiment Basic Writing Score Comparison between Experimental Class and Control Class (Full Score 100)

Class	Number	Average Score	Highest Score	Lowest Score	Standard Deviation
Experimental Class	48	62.35	78	45	6.12
Control Class	48	62.18	77	44	6.08

TABLE 2. Pre- and Post-Experiment Average Writing Score Change Comparison between the Two Classes (Full Score 100)

Class	Pre-Experiment Average Score	Post-Experiment Average Score	Score Improvement	Improvement Rate
Experimental Class	62.35	75.68	13.33	21.38%
Control Class	62.18	68.42	6.24	10.04%

TABLE 3. Post-Experiment Average Number of Various Grammatical Errors per Student in the Two Classes

Error Type	Average Errors in Experimental Class	Average Errors in Control Class	Error Reduction Range
Word Spelling Errors	2.16	4.85	2.69
Tense and Voice Errors	2.08	5.32	3.24
Subject-Verb Collocation Errors	1.53	3.96	2.43
Sentence Pattern Hybrid Errors	1.86	4.28	2.42
Discourse Logic & Argumentation Errors	1.12	3.15	2.03

TABLE 4. Statistics of Students' Autonomous Learning Behaviors in Writing (Per Person per Semester)

Statistical Item	Experimental Class	Control Class	Difference
Autonomous Composition Revision Times	8.65	3.22	5.43
Active Reference to Model Essays	6.38	2.15	4.23
Active Sorting of Writing Knowledge Points	7.52	2.86	4.66

using the system, ignoring guidance on students' writing thinking, emotional expression, and innovative conception.

(4) Insufficient deep correction ability for Chinglish: The system can identify obvious Chinglish grammatical errors but needs to strengthen the recognition and optimization of implicit Chinglish logic and sentence problems caused by cultural and thinking differences.

B. OPTIMIZATION STRATEGIES AT THE ALGORITHMIC AND TECHNICAL LEVEL

(1) Upgrade pre-trained language models, expand the English writing corpus for Chinese students, and improve the recognition accuracy of complex semantics and implicit errors.

(2) Optimize multi-algorithm integration logic, increase the evaluation weight of discourse conception, argumentation logic, and emotional expression, and improve the feedback function of high-level writing dimensions.

(3) Set a hierarchical feedback mode to match revision suggestions and learning materials of different difficulties for students at basic, intermediate, and advanced levels.

C. OPTIMIZATION STRATEGIES AT THE TEACHING APPLICATION LEVEL

(1) Establish a mixed correction mode of "intelligent system + manual teaching". The system is responsible for automatic feedback on basic grammar, vocabulary, and sentence patterns, while teachers focus on commenting on article conception, logical thinking, and innovative expression, complementing each other.

(2) Teachers carry out targeted classroom teaching based on system learning status statistics, focusing on explaining high-frequency error knowledge points of the class.

(3) Guide students to use the automatic feedback system correctly, emphasize that the system is an auxiliary learning tool, encourage independent thinking, and cultivate autonomous writing thinking.

D. OPTIMIZATION STRATEGIES AT THE PLATFORM OPERATION LEVEL

(1) Regularly update composition question banks, model essay materials, and grammar rule libraries to keep pace with teaching progress and examination trends.

(2) Simplify the system operation interface, adapt to multi-terminal use of mobile phones and computers, and lower the use threshold for students and teachers.

(3) Add expanded learning resources such as writing knowledge micro-lectures, sentence pattern templates, and common collocations to make the feedback system have both correction and teaching functions.

VI. CONCLUSION AND PROSPECT

A. RESEARCH CONCLUSIONS

Supported by NLP algorithms, this study integrates core technologies such as text segmentation, part-of-speech tagging, grammatical error correction, semantic analysis, and

intelligent scoring to construct an automatic feedback model for English writing. Through a one-semester controlled experiment and four sets of statistical tables, this paper quantitatively analyzes the application effect of the system. The core conclusions are as follows:

(1) Integrating multiple NLP algorithms can realize full-dimensional automatic detection and real-time feedback of English compositions, covering four core dimensions: vocabulary, grammar, sentence pattern, and discourse.

(2) The automatic feedback system for English writing can effectively improve students' writing scores, reduce the number of grammatical errors, and cultivate autonomous writing learning habits.

(3) The intelligent automatic feedback system can reduce teachers' correction workload, provide objective learning status data, and assist teachers in carrying out precise writing teaching.

(4) Current NLP technology still has deficiencies in complex semantic recognition, high-level writing evaluation, and differentiated personalized feedback.

B. RESEARCH DEFICIENCIES

This study has two limitations:

(1) The experimental samples only select non-English majors from one university with a limited scope. Future experiments can be expanded to primary and secondary schools and multiple institutions.

(2) The model focuses on basic error feedback in written English writing and does not conduct in-depth research on the adaptability to oral English writing and academic paper writing.

C. FUTURE RESEARCH PROSPECTS

Future research can be carried out in three directions:

(1) Continuously optimize NLP deep learning models to improve the ability to recognize and feedback complex discourse logic, high-level conception, and implicit Chinglish.

(2) Expand application scenarios to adapt the automatic feedback model to primary and secondary school English writing, IELTS/TOEFL writing, academic English writing, and other scenarios.

(3) Combine big data and artificial intelligence to build an intelligent English writing teaching platform integrating writing training, intelligent correction, learning status analysis, and personalized teaching recommendation, comprehensively promoting the intelligent development of English education.

AUTHOR CONTRIBUTIONS

Conceptualization – Yong Liu and Xiaoliang Shi; methodology – Yong Liu and Xiaoliang Shi; investigation – Yong Liu and Xiaoliang Shi; writing-original draft preparation – Yong Liu and Xiaoliang Shi. All authors have read and agreed to the published version of the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

FUNDING

This research was supported by, Project supported by Scientific Research Found of Major Theoretical and Practical Issues Research Project of Philosophy and Social Sciences of Shaanxi Province 2022 International Communication Capacity Building Key Project (Cooperation Project) (Project No.:2022HZ0866)

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] L. Yang, L. J. Zhang, and H. R. Dixon, "Understanding the impact of teacher feedback on EFL students use of self-regulated writing strategies," Journal of Second Language Writing, vol. 60, Art. no. 101015, 2023, doi: 10.1016/j.jslw.2023.101015.

[2] H. Alisoy, "The Role of Teacher Feedback in Enhancing ESL Learners Writing Proficiency," Global Spectrum of Research and Humanities, vol. 1, 2024, doi: 10.69760/1rmdzx45.

[3] Y. Wang, et al., "Chinese EFL Teachers Writing Assessment Feedback Literacy: A Scale Development and Validation Study," Assessing Writing, vol. 56, Art. no. 100726, 2023, doi: 10.1016/j.asw.2023.100726.

[4] A. Charalampous and D. Maria, "The Contribution of Teacher Feedback to the Revision of Students Work in Primary and Secondary Education: A Systematic Literature Review," International Journal of Learning and Development, vol. 14, pp. 18, 2024, doi: 10.5296/ijld.v14i3.21991.

[5] C. Omoeva, N. Menezes Cunha, and W. Moussa, "Measuring equity of education resource allocation: An output-based approach," International Journal of Educational Development, vol. 87, Art. no. 102492, 2021, doi: 10.1016/j.ijedudev.2021.102492.

[6] L. Yang and S. Zhao, "ARGUS: A Neuro-Symbolic System Integrating GNNs and LLMs for Actionable Feedback on English Argumentative Writing," Systems, vol. 13, no. 12, 2025, doi: 10.3390/systems13121079.

[7] Supriyono et al., "Advancements in natural language processing: Implications, challenges, and future directions," Telematics and Informatics Reports, vol. 16, Art. no. 100173, 2024, doi: 10.1016/j.teler.2024.100173.

[8] Y. Xu, J. Zhong, and Y. Han, "How teacher feedback literacy changed by using Generative Artificial Intelligence (GenAI) feedback as exemplars: Case studies of English as a foreign language school teachers in China," Teaching and Teacher Education, vol. 176, Art. no. 105501, 2026, doi: 10.1016/j.tate.2026.105501.

[9] W. Wiboolyasarin, et al., "Synergizing collaborative writing and AI feedback: An investigation into enhancing L2 writing proficiency in wiki-based environments," Computers and Education: Artificial Intelligence, vol. 6, Art. no. 100228, 2024, doi: 10.1016/j.caeai.2024.100228.

[10] C. Keh, "Feedback in the writing process: A model and methods for implementation," English Language Teaching Journal, vol. 44, pp. 294-304, 1990, doi: 10.1093/elt/44.4.294.

[11] J. Marie Lim and C. Go, "Integrating ChatGPT into Teacher Feedback: Practical Insights for L2 Writing Instruction," TESL-EJ, vol. 29, no. 3, 2025, doi: 10.55593/ej.29115int.

[12] Y. Mao, et al., "Incorporating emotion for response generation in multi-turn dialogues," Applied Intelligence, vol. 52, no. 7, pp. 7218-7229, 2022, doi: 10.1007/s10489-021-02819-z.

[13] I. Saputra, et al., "The Evolution of Educational Assessment: How Artificial Intelligence is Shaping the Trends and Future of Learning Evaluation," The Indonesian Journal of Computer Science, vol. 13, 2024, doi: 10.33022/ijcs.v13i6.4465.

[14] L. Tian and Y. Zhou, "Learner engagement with automated feedback, peer feedback and teacher feedback in an online EFL writing context," System, vol. 91, Art. no. 102247, 2020, doi: 10.1016/j.system.2020.102247.

[15] C. Kepner Goring, "An Experiment in the Relationship of Types of Written Feedback to the Development of Second-Language Writing Skills," The Modern Language Journal, vol. 75, no. 3, pp. 305-313, 2011, doi: 10.1111/j.1540-4781.1991.tb05359.x.

[16] K. L. Partridge, "A comparison of the effectiveness of peer vs. teacher evaluation for helping students of English as a second language to improve the quality of their written compositions," 1981.

[17] T. M. Paulus, "The effect of peer and teacher feedback on student writing," Journal of Second Language Writing, vol. 8, no. 3, pp. 265-289, 1999, doi: 10.1016/S1060-3743(99)80117-9.

[18] U. Connor and K. Asenavage, "Peer response groups in ESL writing classes: How much impact on revision?" Journal of Second Language Writing, vol. 3, no. 3, pp. 257-276, 1994, doi: 10.1016/1060-3743(94)90019-1.

[19] Y. Attali, "Exploring the Feedback and Revision Features of Criterion," Journal of Second Language Writing, vol. 14, 2004.

[20] L. Kohnke, B. L. Moorhouse, and D. Zou, "Exploring generative artificial intelligence preparedness among university language instructors: A case study," Computers and Education: Artificial Intelligence, vol. 5, Art. no. 100156, 2023, doi: 10.1016/j.caeai.2023.100156.

[21] Y. Su, Y. Lin, and C. Lai, "Collaborating with ChatGPT in argumentative writing classrooms," Assessing Writing, vol. 57, Art. no. 100752, 2023, doi: 10.1016/j.asw.2023.100752.