Hindawi
Mobile Information Systems
Volume 2022, Article ID 5073924, 12 pages
https://doi.org/10.1155/2022/5073924

*Research Article*

# Survey on the Current Situation of College English Teaching in China's Universities and the Direction of College English Teaching Reform and Development

**Ruifang Duan and Peilei Qiu**

*School of Foreign Languages, Zhengzhou Business University, Zhengzhou, Henan 451200, China*

Correspondence should be addressed to Peilei Qiu; georgeqiu@zbu.edu.cn

Received 27 June 2022; Revised 8 August 2022; Accepted 31 August 2022; Published 20 September 2022

Academic Editor: Yajuan Tang



With the development of internationalization, English learning becomes more and more important. In the process of language learning, writing has always played a very important role. A writer's language proficiency can be improved by the amount of reading experience and knowledge, which is necessary to produce high-quality writing. In recent years, there have been many writing assistant recommendation systems supported by different technical means, which provide great help for college students' writing. In order to solve the problem that traditional recommendation algorithms can not recommend accurately, this paper proposes a hybrid recommendation algorithm and applies it to the recommendation of English writing documents. The algorithm generates three-dimensional feature vectors by learning the characteristics of students like, dislike, and similar students. Three low-dimensional feature vectors are linearly combined to form the representation vector of college students. And the cosine similarity is used as the similarity index to recommend English writing literature related to similar college students to the target college students, so as to achieve the recommendation of English writing literature. Experimental results show that this recommendation algorithm is superior to the other four algorithms in mean absolute error (MAE) and time performance and has high recommendation quality.

## 1. Introduction

English teaching is to develop students' international perspective and make them build up the willingness to learn independently and consciously [1]. Learning to write not only strengthens students' thinking skills and improves their expressive abilities, but also measures the effectiveness of teachers' teaching [2].

Since 2002, research on college English writing instruction has grown rapidly and still lacks a theoretical foundation in terms of the single research topic previously studied [3]. From 2000 to 2009, research has shown an overall upward trend. Specifically, it began in 2002, while research and development has declined since 2010. Over the past decade, Chinese researchers have explored and learned from their research experiences. Foreign scholars have shown a differentiated development in their research on university English language teaching by combining it with other areas of foreign language studies [4]. The research themes are now very rich and the theoretical backgrounds vary. Researchers have been more concerned with issues related to the teaching of college English writing, which provides a solid foundation for empirical research [5, 6]. The following are specific studies by different scholars at home and abroad on improving English writing (see Tables 1 and 2).

A recommendation system is an important tool to help users deal with information overload in the era of big data. It identifies a set of items of interest based on user behavior and recommends items of interest to users, saving them time.

Collaborative filtering (CF) algorithm is known as one of the successful techniques for personalized recommendation systems. The collaborative filtering technique recommends items for a target user by identifying content to be recommended by other users with similar interests. Collaborative

Table 1: Domestic literature review details on English writing.

| Literature | Author | Year | Methodological characteristics |
| --- | --- | --- | --- |
| [7] | Lin et al. | 2018 | The effect of cooperative learning on non-English majors' English writing by taking questionnaires and interviews with non-English majors. The results found that the use of cooperative learning strategies significantly reduced the overall anxiety, somatic anxiety, and evaluation anxiety of English writing among non-English major college students. |
| [8] | Wang | 2020 | A classroom experiment in college English classrooms to investigate the effects of collaborative writing on the second language development of college English learners in China. |
| [9] | Zhou | 2020 | An experiment to explore the help of Chinese and American college students' English writing skills through online communication and tutoring. The results showed that this approach based on online tutoring and communication between Chinese and American college students was very helpful in improving Chinese college students' writing skills. |
| [10] | Chen et al. | 2019 | Implemented a comparative teaching study of cross-cultural e-mail communication and traditional paper-and-pencil writing teaching methods, respectively. It showed that cross-cultural e-mail communication significantly improved students' writing skills, especially fluency and accuracy. |
| [11] | Tan | 2019 | A series of studies on students' motivation and ability to write, the impact of college students' writing, and writing instruction through a model based on digital writing instruction. |

Table 2: Foreign literature review details on English writing.

| Literature | Author | Year | Methodological characteristics |
| --- | --- | --- | --- |
| [12] | Noorizadeh-Honami et al. | 2018 | Second language writing is a complex phenomenon in which emotional factors play an important role. Among them, the level of motivation determines whether students can successfully organize the complex writing process and the quality of their second language writing. |
| [13] | Graham S. | 2019 | Investigated the English writing skills of high school students in two countries, Germany and Switzerland, two years before and one year after graduation. It looks at the level and development of English writing, as well as differences between groups (country, gender, and language background). |
| [14] | Teng L. S. | 2018 | Draws on the requirements for writing research as a lingua franca. Shifting the focus from multilingual speakers to proficient English speakers. |
| [15] | Shafqat A. | 2020 | Presented a computational analysis of English writing and developed new structures to describe this relationship between spelling and sound. |
| [16] | Canagarajah S. | 2018 | Identified all traditional, textbook idioms in the British Corpus of Spoken (Basic) academic English, examining the range of texts in which they occur. Their frequencies were then compared to the same idioms used in the Oxford Corpus of Academic English (OCAE). |

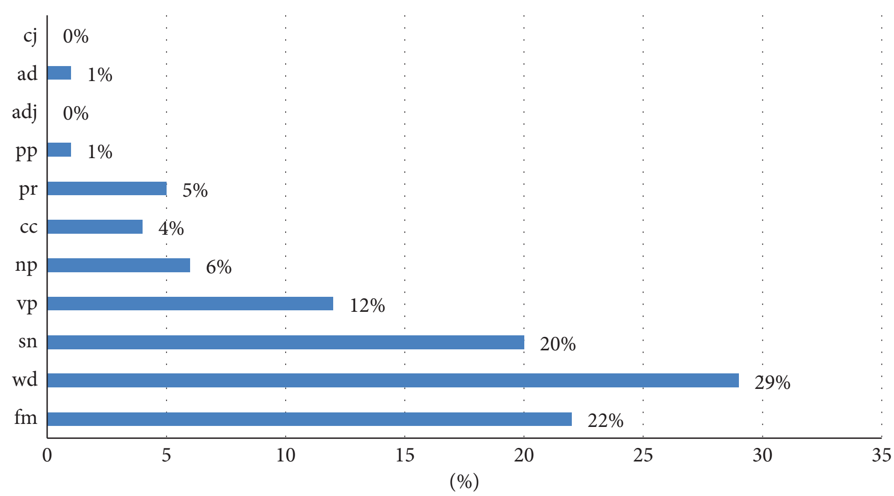
filtering recommendation systems have been implemented in different application areas. News: GroupLens system uses collaborative filtering to help users find the right content for their needs from a large database of news. Social: Ringo is an online social information filtering system that uses collaborative filtering to build user profiles based on their ratings in music albums. Third, the e-commerce domain: Amazon uses topic diversification algorithms to improve its recommendations. The system uses collaborative filtering techniques to generate tables of similar items offline from an item-to-item matrix as a way to overcome scalability issues.

With the computerization of university management system, the recommendation system is also used for university teaching. English learners spend a lot of effort on English learning, but in terms of the results of learning, they do not achieve the expected results [17]. The biggest reason is because students lack a lot of reading experience and knowledge base [18]. The main way to improve college students' English writing skills is through reading a lot of excellent English writing literature [19]. How to meet the needs of students with different English levels, provide accurate and personalized bibliography for each user among the huge amount of reference literature, and make real-time recommendations through an online recommendation system is the purpose of this research paper. Therefore, this paper proposes a hybrid recommendation algorithm based on multidimensional feature representation learning (MFL). The algorithm splits English writing literature scoring network. The algorithm based on the improved LINE performs hierarchical advancing learning of students' favorite English writing literature and aversion to English writing literature. Based on the improved DeepWalk algorithm, a sequence of similar students is obtained and similar student features are captured. The preferred features, disliked features, and similar student features are linearly combined and connected as the final feature vector of students. The cosine similarity is used as the similarity metric to complete the English writing literature recommendation task.

The innovative points of this paper are as follows:

Table 3: Overview of the distribution of writing errors among non-English major learners (percentages are retained to two decimal places).

| Dimensions | G1 | G2 | Total | Percentage (%) |
|---|---|---|---|---|
| fm | 375 | 406 | 781 | 22.97 |
| wd | 505 | 520 | 1025 | 30.15 |
| sn | 342 | 199 | 541 | 15.91 |
| vp | 213 | 230 | 443 | 13.03 |
| np | 100 | 102 | 202 | 5.94 |
| cc | 64 | 108 | 172 | 5.06 |
| pr | 96 | 43 | 139 | 4.09 |
| pp | 12 | 45 | 57 | 1.68 |
| aj | 1 | 0 | 1 | 0.03 |
| ad | 17 | 22 | 39 | 1.15 |
| cj | 0 | 0 | 0 | 0.00 |
| Total | 1725 | 1675 | 3400 | |
| Percentage (%) | 50.74 | 49.26 | | |




| | fm | wd | sn | vp | np | cc | pr | pp | adj | ad | cj |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ■ | 22% | 29% | 20% | 12% | 6% | 4% | 5% | 1% | 0% | 1% | 0% |

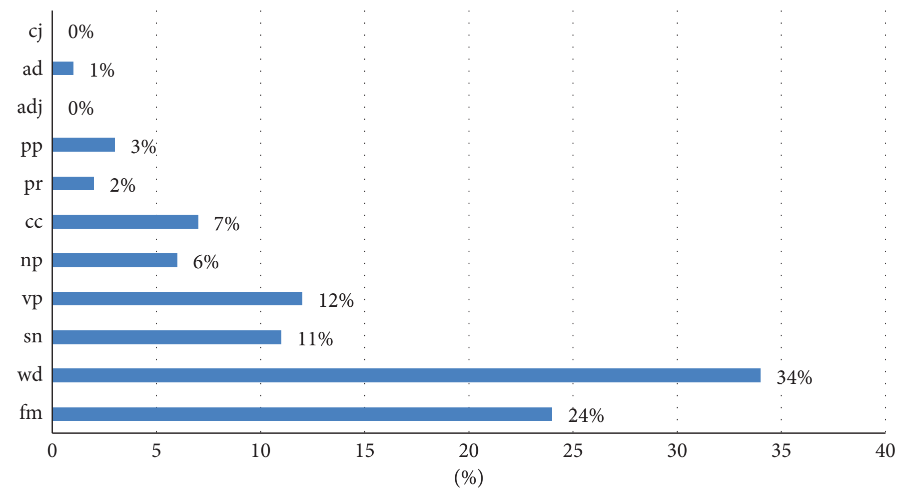
Figure 1: Distribution of English writing errors among freshmen.

(1) The English writing literature scoring network was split to provide a hierarchical advancement of students' favorite and averse English writing literature based on an improved LINE algorithm.
(2) Based on the improved DeepWalk algorithm, we obtain similar student sequences and capture similar student features.
(3) After linear combination of the preferred features, aversive features, and similar student features, they are connected as the final feature vector of students, and the cosine similarity is used as the similarity metric to complete the English writing literature recommendation task.

This paper consists of four main parts: the first part is the introduction, the second part is methodology, the third part is result analysis and discussion, and the fourth part is the conclusion.

## 2. Methodology

*2.1. Survey on the Current Situation of College English Teaching in China's Universities.* Based on the error classification method of [20] in CLEC, a total of 11 categories of errors were classified into language errors. They are word form errors-fm, lexical errors-wd, syntactic errors-sn, verb phrase errors-vp, noun phrase errors-np, collocation errors-cc, pronoun errors-pr, preposition errors-pp, adjective errors-adj, adverb errors-ad, and conjunctive errors-cj. After labeling the student essay samples, the writing errors of the college student essay samples in both grades were retrieved and counted separately. The amount of writing errors in the freshman composition was labeled as $G_1$ and the amount of




| | fm | wd | sn | vp | np | cc | pr | pp | adj | ad | cj |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ■ | 24% | 34% | 11% | 12% | 6% | 7% | 2% | 3% | 0% | 1% | 0% |

FIGURE 2: Distribution of English writing errors among sophomores.

writing errors in the sophomore composition was recorded as $G_2$, and the results are shown in Table 3.

The data in Table 3 are based on the number of writing errors in English compositions of freshmen and sophomores. In order to clarify the distribution of writing errors of college students more clearly and intuitively, a pie chart is used to analyze it (see Figures 1 and 2).

From the above table and graphs, it can be seen that the most frequent errors in students' writing are vocabulary errors (wd). The results are 505 vocabulary errors for freshmen and 520 vocabulary errors for sophomores. The total number of errors in vocabulary (wd) for both grades was 1025, accounting for 30.15% of all errors in both grades. Students made the second highest number of errors in word form (fm), with 375 errors for freshmen and 406 errors for sophomores. The total number of errors in word form (fm) for students in both grades was 781, accounting for 23.27% of all errors in both grades. Syntactic (sn) errors were also very high for both grades, with 342 errors for freshmen and 199 errors for sophomores. The total number of syntactic (sn) errors for both grades was 15.91% of all errors for both grades, which was the third highest.

The errors of verb phrases (vp), noun phrases (np), collocations (cc), and pronouns (pr) in the two grades ranked fourth, fifth, sixth, and seventh. Among them, the number of verb phrase (vp) errors of freshmen is 213. The number of verb phrase errors of sophomores is 230. There are 443 errors in verb phrases (vp) of students in the two grades, accounting for 13.03% of all errors in the two grades. The number of NP errors of freshmen is 100. The number of noun phrase errors of sophomores is 102. There are 202 errors in noun phrases (np) made by students in the two grades, accounting for 5.94% of all errors in the two grades. The number of errors in collocation (cc) of freshmen is 64. The number of collocation errors of sophomores is 108. There are 172 errors in collocation (cc) in the two grades, accounting for 5.06% of all errors in the two grades. The number of errors in pronouns (pr) of freshmen is 96. The number of pronoun errors of sophomores is 43. There are 139 pronoun (pr) errors in the two grades, accounting for 4.09% of all errors in the two grades. The number of errors in preposition (pp) of freshmen is 12. The number of preposition errors of sophomores is 45. There are 57 errors in preposition (pp) made by students in the two grades, accounting for 1.68% of all errors in the two grades, and the amount of errors ranks eighth. The adverb (ad) ranks ninth in the error quantity. The number of errors in adverbs (ad) of freshmen is 17. The number of adverb errors of sophomores is 22. The students in the two grades made a total of 39 errors in adverbs (ad), accounting for 1.15% of all errors in the two grades. The total number of errors in adjectives (adj) in the two grades is 1, accounting for 0.03% of all errors in the two grades, and the number of errors ranks tenth. In addition, no errors in conjunctions (cj) were found in both grades, and no errors in adjectives (adj) were found in sophomores.

In addition, freshmen made a total of 1725 errors in 11 dimensions, accounting for 50.73% of all errors. Sophomores made a total of 1675 errors in these 11 dimensions, accounting for 49.26% of all errors.

The two graphs above compare the differences in the amount of writing errors between the two grades in two different ways. The bar chart in Figure 3 and the line graph in Figure 4 clearly show that sophomores made more errors than freshmen in seven dimensions: fm, wd, vp, np, cc, pp, and ad. However, in the three dimensions sn, pr, and adj, freshmen have more errors than sophomores. Among them, the amount of errors of freshmen students on sn was 149 higher than that of sophomores; the amount of errors of freshmen students on pr was 44 higher than that of sophomores; and the amount of errors of freshmen students on adj was 3.0 higher than that of sophomores.

In order to investigate the differences in writing errors among freshmen and sophomores at different levels, the top

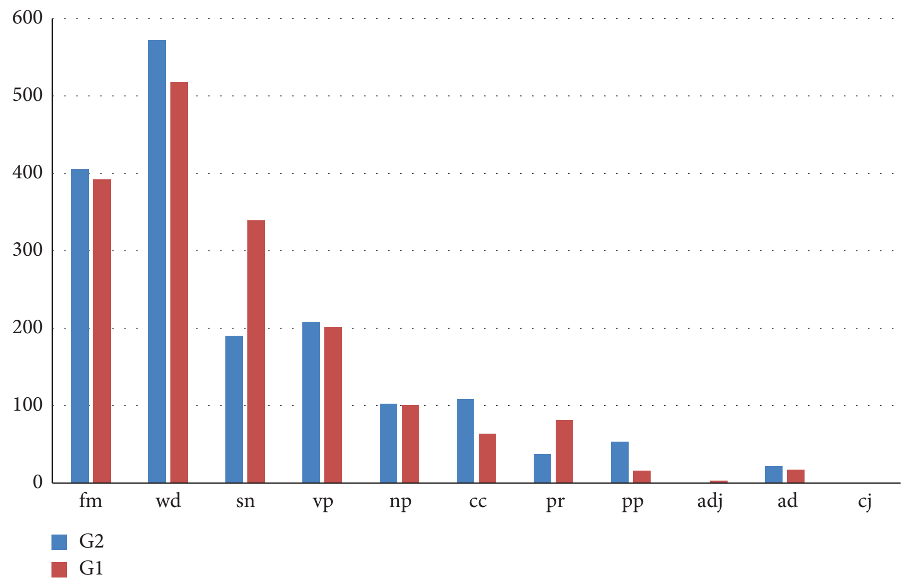



Figure 3: Bar chart comparing the difference in writing errors between the two grades.

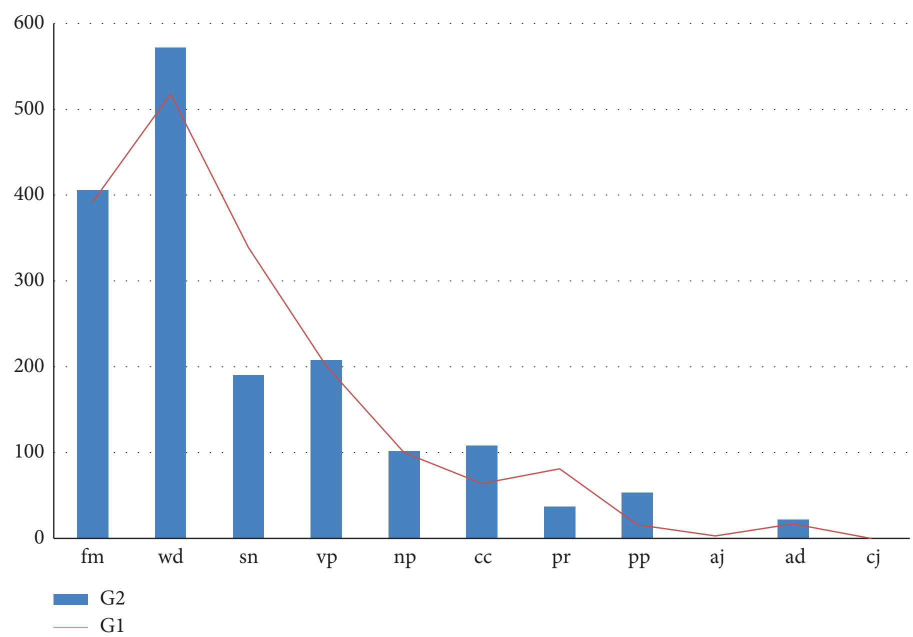



Figure 4: Line graph comparing the variability of writing errors between the two grades.

30 students in each grade were classified as the high group and the last 30 students were classified as the low group. The analysis of the writing error data revealed that the distribution of the data met the normal distribution. Therefore, an independent sample S-test was conducted to investigate the differences in writing errors between the high and low subgroups of freshmen and sophomores (see Tables 4 and 5).

Table 4 shows that freshmen made zero errors on the conjunction (cj), while they made errors on the other ten dimensions. In addition, the means of the lower subgroups of freshmen were greater than the means of the higher subgroups on all of these dimensions. This indicates that the low subgroup of freshmen made more errors on the ten dimensions fm, wd, sn, vp, np, cc, pr, pp, adj, and ad than the high subgroup.

From the results of the independent samples S-test in Table 4, there is a significant difference between the freshman high and low subgroups on the seven dimensions of fm, wd, sn, vp, np, cc, and pr (Sig. (two-sided) $< 0.05$). This indicates that the number of errors in the freshman low group was much higher than that in the freshman high group in these seven dimensions. This also indicates that the amount of errors in the freshman low group is not significantly different from that in the freshman high group in these three dimensions.

The data in Table 5 show that the sophomores had zero errors in adjectives (adj) and conjunctions (cj) and some errors in all nine dimensions: fm, wd, sn, vp, np, cc, pr, pp, and ad. In addition, the means of the sophomore low subgroup

Table 4: Analysis of the variability of each dimension in the high and low subgroups of freshman year.

| Dimensions | High group | | Low group | | Independent samples S-test | |
|---|---|---|---|---|---|---|
| | Mean value | Standard deviation | Mean value | Standard deviation | *S*-value | *S*-test |
| fm | 0.95 | 0.854 | 5.25 | 2.358 | −10432 | 0 |
| wd | 2.31 | 1.013 | 6.94 | 2.463 | −12.427 | 0 |
| sn | 1.08 | 0.924 | 4.54 | 2.792 | −10.579 | 0 |
| vp | 0.39 | 0.287 | 2.97 | 0.975 | −5.653 | 0 |
| np | 0.58 | 0.968 | 2.19 | 1.363 | −6.815 | 0 |
| cc | 0.29 | 0.231 | 0.85 | 0.473 | −2.294 | 0 |
| pr | 0.36 | 0.574 | 1.08 | 0.995 | −3.718 | 0 |
| pp | 0.11 | 0.61 | 0.17 | 0.868 | −1.593 | 0.106 |
| adj | 0 | 0 | 0.08 | 0.289 | −1.672 | 0.104 |
| ad | 0.13 | 2.914 | 0.17 | 0.373 | −0.534 | 0.605 |
| cj | 0 | 0 | 0 | 0 | none | none |

Table 5: Analysis of the variability of each dimension in the high and low subgroups of the sophomores.

| Dimensions | High group | | Low group | | Independent samples S-test | |
|---|---|---|---|---|---|---|
| | Mean value | Standard deviation | Mean value | Standard deviation | *S*-value | Sig. (two-sided) |
| fm | 1.09 | 0.857 | 6.56 | 2.688 | −9.163 | 0 |
| wd | 3.08 | 1.589 | 7.7 | 2.374 | −12.995 | 0 |
| sn | 0.82 | 1.326 | 2.76 | 2.482 | −9.619 | 0 |
| vp | 0.92 | 1.132 | 2.94 | 2.086 | −6.357 | 0 |
| np | 0.47 | 0.704 | 1.44 | 1.139 | −7.592 | 0 |
| cc | 0.51 | 0.895 | 1.53 | 1.364 | −7.668 | 0 |
| pr | 0.12 | 0.879 | 0.58 | 0.896 | −6.542 | 0 |
| pp | 0.18 | 0.39 | 1.03 | 1.008 | −2.492 | 0 |
| adj | 0 | 0 | 0 | 0 | none | none |
| ad | 0 | 0 | 0.4 | 1.152 | −0.938 | 0.085 |
| cj | 0 | 0 | 0 | 0 | none | none |

Table 6: Correlation analysis between the amount of writing errors and essay scores in the first year of college.

| | | Composition score | Number of errors |
|---|---|---|---|
| Composition score | Correlation coefficient | 1 | $-0.967^{**}$ |
| | Significance (two-tailed) | | 0 |
| | *T* | 115 | 115 |
| Number of errors | Correlation coefficient | $-0.967^{**}$ | 1 |
| | Significance (two-tailed) | 0 | |
| | *T* | 115 | 115 |

students were greater than the means of the sophomore high subgroup on all nine dimensions. It means that, in all nine dimensions, the amount of errors of the sophomore low subgroup is more than that of the sophomore high subgroup.

From the results of the independent samples S-test in Table 5, there were significant differences (Sig. (two-sided) $< 0.05$) between the sophomore high and low subgroups of students on the eight dimensions of fm, wd, sn, vp, np, pr, cc, and pp. This indicates that the amount of errors in these eight dimensions was much higher in the sophomore low group than in the freshman high group. In other words, the amount of errors of the sophomore low group was not significantly different from that of the sophomore high group on the dimension of adverb (ad).

The Pearson correlation analysis leads to the data in Table 6. The specific explanation is as follows: the amount of writing errors in the freshman students' composition sample was negatively correlated with the students' composition scores, with a correlation coefficient $|r|$ of −0.967; i.e., their correlation was extremely high. From the above, it is clear that the higher the amount of writing errors in the English composition sample of freshmen and sophomores, the lower their composition scores.

*2.2. English Writing Recommendation Algorithm.* The MFL recommendation algorithm has four steps.

(1) The matrix of students' ratings of English writing documents is considered as a complex network, where students and English writing documents are considered as network nodes and ratings are considered as network linkage weights. Using the linkage weights as

a distinction, the network is divided into high-weight subnetworks and low-weight subnetworks.

(2) Based on the improved LINE algorithm, the network structure of the high-weight subnetwork is learned and the student vector and the English writing literature vector are generated. The English writing literature vector generated by the high-weight subnetwork is used as the input of the low-weight network learning to learn the structure of the low-weight subnetwork and generate the student vector of the low-weight subnetwork.

(3) From the whole network of nodes, student nodes with the same rating on English writing literature are randomly selected to form a sequence of student nodes. The sequence of student nodes is fed into the CBOW (continuous bag-of-words) algorithm to learn the features of similar students.

(4) The three-dimensional feature vectors generated by each student node are linearly combined and stitched together to form the final student vector. The cosine distance of the vector is used as the similarity index between the nodes to generate the set of similar students of the target students. The English writing literature associated with similar students is recommended to the target students to complete the recommendation task.

As the input of the recommendation algorithm, the scoring matrix is usually composed of $\langle p, x, n\rangle$ as records, where $p$ represents the student number; $x$ represents the English writing literature number; and $n$ represents the student's rating of the English writing literature. The students and the English writing literature constitute the nodes in the network. $P = \{p_1, p_2, \cdots, p_w\}$ is the set of students; $X = \{x_1, x_2, \cdots, x_t\}$ is the set of English writing documents; $E = \{e_{xy} \mid x = 1, 2, \cdots, t; y = 1, 2, \cdots, w\}$ is the set of connected edges. When there is a link $e_{xy} = n$, it represents student $p_x$'s score $n$ on English writing literature $x_y$. The network can be expressed as $A = (Q, E, M)$, where $Q = P \cap X$, $E \subset P \times X$, and $M$ is the weight matrix on the connected edges. The network representation learning algorithm learns the network structure information and generates a low-dimensional vector representation of the network nodes. The similarity of the vectors is used as an indicator of student similarity, and the most similar Top-k students are selected. The English writing literature associated with the similar student set constitutes English writing literature recommendation set and is recommended to the target students.

The edges with more than half of the maximum weight of the connected edges of the network are extracted to generate the student favorite network (high-weight subnetwork). For example, for a rating network with a maximum rating of 5, any contiguous edges with a rating greater than or equal to 3 are extracted.

The English writing literature node vector is denoted by $p_y' \in \mathrm{R}^d$ and the student node vector is denoted by $p_x \in \mathrm{R}^d$.

For each edge <*x, y*>, using the Softmax function, the conditional probability of student node $q_x$ generating English writing literature node $q_y$ is as follows:

$$U\left(q_y \mid q_x\right) = \frac{\exp\left(p_y^{'N} \cdot p_x\right)}{\sum_{z=1}^{|Q|} \exp\left(p_z^{'N} \cdot p_x\right)}, \tag{1}$$

where $|Q|$ represents all nodes. There is a strong correlation between $q_y$ and $p_x$, and the English writing literature node ($q_y$) grows exponentially with the student node vector ($p_x$). The empirical distribution of student nodes $q_x$ generating English writing literature nodes $q_y$ is shown in

$$\widehat{U}\left(q_y \mid q_x\right) = \frac{m_{x,y}}{d_x} = \frac{m_{x,y}}{\sum_{z \in z(x)} m_{x,z}}, \tag{2}$$

where $m_{x,y}$ are the weights of the connected edges; $d_x$ is the degree of node $q_x$; and $z(x)$ is the node adjacent to node $q_x$. KL scatter is a function that measures the difference between two probability distributions. In this paper, KL is used to denote the degree of difference between the conditional probability $u(\cdot \mid q_x)$ and the empirical distribution $\widehat{u}(\cdot \mid q_x)$. When the two distributions are the same, the KL scatter is zero, and the greater the difference between the two, the greater the KL scatter. Using the KL scatter, the loss function can be obtained.

$$O = \sum_{x \in Q} \lambda_x D_{KL}\left(U(\cdot \mid q_x) \parallel \widehat{U}(\cdot \mid q_x)\right), \tag{3}$$

where $\lambda_x$ denotes the importance of node $q_x$. Here take $\lambda_x = d_x$, and finally get the objective function as follows:

$$O = -\sum_{(x,y) \in E} m_{xy} lg U\left(q_y \mid q_x\right). \tag{4}$$

In the process of loss optimization, the calculation of the conditional probability $U(q_y \mid q_x)$ requires traversal of the entire network of nodes. In large-scale network structures, the process of calculation requires a lot of time and resources. To solve the problem of higher complexity, the traversal of the whole network nodes is replaced by the negative sampling method, and the loss function is transformed as follows:

$$O = lg\sigma\left(p_y^{'N} \cdot p_x\right) + \sum_{x=1}^{Z} E_{q_t U_t}\left[\lg\sigma\left(p_t^{'N} \cdot p_x\right)\right], \tag{5}$$

where $p_y^{'\mathrm{N}}$ and $p_t^{'\mathrm{N}}$ denote the English writing literature vector; $p_x$ denotes the student vector; $(i) = (11 + \exp\ (-i))$; $Z$ is the number of negative samples; $Z = 5$ is generally chosen for large networks; and $U_t$ is the noise distribution. $U_t(q) \propto d_q^{3/4}$, where $d_q$ represents the degree of English writing literature nodes.

After iterative optimization, two types of vectors are generated. $\Phi_1$ and $\Phi'$ represent the vector of students' preferred features and the vector of English writing literature, respectively.

Table 7: Algorithm 1.

| Algorithm 1. Learning preference and aversion features |
|---|
| Input: Graph A(Q,E,M) |
| Dimensionality $d = 200$ |
| Number of negative samples $Z = 5$ |
| Output: Students' preference features $\varphi 1 \in R\|Q1\| * d$ |
| Students' dislike features in $\varphi 2 \in R\|Q1\| * d$ |
| 1:Initialization in $\varphi 1 \in R\|Q1\| * d$, Q1∊student node |
| $\varphi 2 \in R\|Q1\| * d$ in initialization, Q1∊student node |
| Initialization in $\varphi' \in R\|Q1\| * d$, Q2∊English writing literature node |
| where:Q1∪Q2 = $Q$ |
| 2:Extract the information of network edges $E1 \cup E2 = E$ |
| Where: the weight of edge E1 is greater than 1/2 of the maximum weight |
| The weight of edge E2 is less than 1/2 of the maximum weight |
| 3:According to the degree of the network English writing literature nodes, the negative sampling table of English writing literature is generated |
| Preferred feature table N1, disliked feature table N2 |
| 4 for Exy in E1 do |
| 5: Sampling $Z$ English writing documents from table N1, sampling English writing documents in $\varphi'$ to form a list |
| 6: Input the list, $\varphi 1(x)$ and $\varphi'(y)$ to the objective function |
| 7:end for |
| 8:Get the vector representation $\varphi'$ of the English writing literature nodes |
| 9 for $Exy$ in E2 do |
| 10:sample $Z$ English writing documents from table N2, sample English writing documents $\varphi'$ to form a list |
| 11: Input the list, $\varphi 2(x)$ and $\varphi'(y)$ to the objective function |
| 12. end for |
| 13: return,$\varphi 1 \varphi 2$ |

Extract the edges of the network with even edge weights less than half of the maximum weight to form a student-averse network, and learn the student-averse network with the objective function of learning as shown in

$$O = lg\sigma\left(p_y^{'N} \cdot p_x\right) + \sum_{x=1}^{Z} E_{q_t \sim U_t}\left[\lg\sigma\left(p_t^{'N} \cdot p_x\right)\right], \quad (6)$$

where $p_y^{'N}$ and $p_t^{'N}$ denote English writing literature vectors; $p_x$ denotes the vector of student nodes that need to be relearned; and $Z$, $U_t$ is the same as the setting of preference feature learning to generate the student aversion feature vector $\Phi_2$.

The difference between the aversion network learning and the favorite network learning is mainly reflected in two points. First, in the representation learning of the student favorite network, no initialization settings are made for the student node and the English writing literature node. In the aversion network learning, the English writing literature node needs to be initialized and set. Its setting value is the English writing literature vector output by the favorite network, and the student node is not initialized. Second, during the training process, the preference network learns for all nodes, including the English writing literature node and the student node. During the learning process of the aversion network representation, the English writing literature vector is locked and only the student vector is learned. Algorithm 1 is described in Table 7.

Students with similar characteristics will have similar ratings for the same English writing literature. Inspired by the DeepWalk algorithm, students with the same rating on the same English writing literature are randomly selected. The sequence of randomly selected nodes is treated as statements in natural language processing, and the probability of occurrence of a particular student in a sequence is evaluated as the basis of this part of the algorithm. For a particular English writing literature node $v_j$, students are randomly selected among students who have the same rating on it, forming a sampling sequence $M_{q_y}$. It contains student nodes $M_{q_y}^1, M_{q_y}^2, M_{qy}^3, \cdots, M_{q_y}{}^{\max}$, max being the maximum sequence length. In this paper, the maximum sequence length of randomly selected student sequences is set to max = 100. The actual sampling process may result in the situation that the number of available student nodes for sampling is too small. In order to avoid the resulting repeated training, the minimum value min is set in the sampling process, and the sampling is skipped when the students with the same rating of an English writing literature are less than min. In this paper, min = 10, and multiple groups of similar student sequences constitute a "corpus" for extracting network information.

The extracted sequences are fed into the natural language processing algorithm CBOW model. For a certain English writing document $q_y$, the extracted sequence is $M_{q_y} = (q_1, q_2, q_3, \cdots, q_z)$. This paper hopes to maximize the probability $U(q_x \mid q_1, \cdots, q_{x-1}, q_{x+1} \cdots, q_z)$ by training the corpus. Each of these nodes can be represented by a low-dimensional vector, and subsequently maximizing the probability $U$ can be converted to the following equation:

$$U\left(q_x \mid \Phi(q_1), \cdots, \Phi(q_{x-1}), \Phi(q_{x+1}), \cdots, \Phi(q_z)\right). \quad (7)$$

Table 8: Algorithm 2.

| Algorithm 2. Learning similar student features |
|---|
| Input: Graph A(Q,E,M) |
| Maximum sequence length of sampled users |
| max = 100 |
| Number of iterations num = 30 |
| Minimum sequence length of sampled sequences |
| min = 10 |
| Dimensionality $d = 200$ |
| Window_ size = 40 |
| Output: Student feature vector in $\varphi_3 \in R^{\lvert Q1 \rvert} * ^d$ |
| 1: Initialize $\varphi_3 \in \mathrm{R}^{\lvert Q1 \rvert} * ^{\mathrm{d}}, Q_1$ ∈student node |
| Where: $Q_1 \cup Q_2 = \mathrm{Q}$, $Q_2 \in$English writing literature node |
| 2: for $x = 0$ to num do |
| 3: Randomize the sequence of English writing documents in $Q_2$ to generate the set O |
| 4: for $Q_y \in \mathrm{O}$ do |
| 5: Similar user list = Get Sample (A, $Q_y$, max, min) |
| 6: CBOW($\varphi_3$,list, window_ _size) |
| 7: end for |
| 8: end for |
| 9: return $\varphi_3$ |

In the model of network representation learning, the final optimization objective function is transformed into

$$\min_{\Phi} -lgU\left(q_x \mid \Phi(q_1), \cdots, \Phi(q_{x-1}), \Phi(q_{x+1}) \cdots, \Phi(q_z)\right). \quad (8)$$

In order to reduce the computational effort, the subsequence consisting of the length of a window before and after the target word $q_x$ is selected as the input of the CBOW model during the actual training. In this paper, the window size = 40. The student feature vector $\Phi_3$ is generated through the learning of similar students, and the algorithm is described as shown in Table 8.

For a student node, three sets of vectors $\Phi_1$, $\Phi_2$, $\Phi_3$ will be generated, which represent student preference features, dislike features, and similar student features, respectively. The final low-dimensional vector of student node $q_x$ can be represented as a linear combination of $\Phi_1$, $\Phi_2$, $\Phi_3$:

$$\Phi(q_x) = \left(\alpha_1 \cdot \Phi_1(q_x), \alpha_2 \cdot \Phi_2(q_x), \alpha_3 \cdot \Phi_3(q_x)\right). \quad (9)$$

Based on the experimental experience, $\alpha_1$, $\alpha_2$, and $\alpha_3$ were set to 0.5, 0.3, and 0.2, respectively. The final student vector $\Phi(q_x)$ was generated, and the cosine similarity was used as the interstudent similarity index Sim. The three most similar students were selected to form the similar student set $S(p)$ of the target students.

$$Sim\left(q_x, q_y\right) = \frac{\Phi(q_x) \cdot \Phi\left(q_y\right)}{\| \Phi(q_x) \|_2 \times \| \Phi\left(q_y\right) \|_2}. \quad (10)$$

Recommend all the associated English writing documents of the student set $S(p)$ to the target students to complete the recommendation task.

## 3. Result Analysis and Discussion

*3.1. Suggestions for Improving the Teaching of English Writing in College.* Some suggestions are given here for the causes of the errors in English writing of college students.

(1) Teachers should have the right attitude of error correction and writing guidance. To improve students' English writing, teachers must thoroughly review and instruct students' essays, to correct errors correctly and appropriately. If the error is due to carelessness, it may be caused by exam-like stress or carelessness, or it may be caused by the student's psychological state at the time. In such cases, the student can correct the error on his or her own, or the teacher can correct it by prompting the student. However, errors arising from students' incomplete understanding of the rules of English become an area where teachers must pay attention to helping students master the rules of the target language and correct writing errors.

(2) Teachers can enrich the way they correct essays. There are certain principles that should be followed for essay correction, and a combination of correction modes is more conducive to improving students' writing errors. For example, in practice, a combination of correction methods can be adopted. First of all, common correction symbols can be set, so that students understand the common methods of correction symbols and annotation methods. And the prerequisite for correcting essays is to standardize the use of deletions, additions, adjustments, changes, and other revision marks. Secondly, peer assessment, self-correction, group correction, revision, and teacher correction can all be ways of making corrections. This is about giving students some responsibility in making corrections so that they can speed up the process and mobilize their own subjective awareness.

(3) Improve the teaching methods and strategies of English writing and set writing tasks reasonably. The results-based approach has had a profound impact on the teaching practice of English writing courses in China. The outcomes-based approach has been the "teacher's model explanation-student's imitation-teacher's evaluation" method. Teachers should encourage students to use specific vocabulary and avoid using general, broad vocabulary to better improve the accuracy of English writing and phrasing and to make better use of basic English knowledge.

*3.2. Experiments on English Writing Recommendation Algorithm.* An article bank of English writing containing 100,000 scored data pieces from 943 students on 1,682 English writing documents provided by a university was used for the experiment. The dataset was randomly cut into 80% training set and 20% test set. MAE and root mean square error (RMSE) were used as the measures.

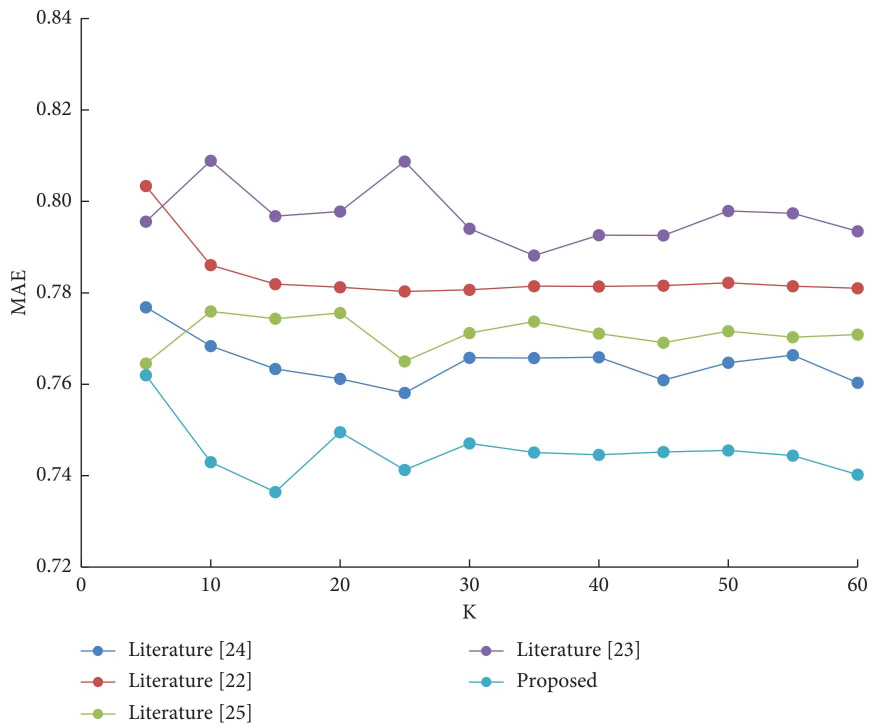



Figure 5: MAE values of 5 algorithms.

Table 9: Running times of the five algorithms at different values of K.

| Algorithm | 50 | 100 | 150 | 200 | 250 |
|---|---|---|---|---|---|
| [21] | 18 | 45 | 82 | 114 | 190 |
| [22] | 181 | 505 | 973 | 1436 | 2164 |
| [23] | 22 | 55 | 100 | 141 | 207 |
| [24] | 196 | 519 | 1000 | 1865 | 2370 |
| Proposed | 29 | 70 | 114 | 152 | 225 |

*Experiment 1.* Comparison of this paper's algorithm with other algorithms' MAE.

The proposed algorithm is compared with the other four algorithms in [21–24]. The recommendation effect of the five algorithms is based on the change of MAE, as shown in Figure 5.

It can be seen that the algorithm of [21] ignores the influence of student attribute characteristics on student trust. And [22] did not introduce the time factor into the similarity calculation, which led to poor recommendation quality. The algorithm in this paper incorporates the improved algorithm of [23] and the algorithm of [24], which significantly improves the recommendation effect and solves the problem of student and project cold start at the same time, so the recommendation accuracy is the highest.

*Experiment 2.* Comparison of the time performance of the proposed algorithm with other algorithms.

In order to better verify the time performance under different data sets, K is selected from 50 to 250, and the value is taken every 50 (see Table 9).

As shown in Table 9, [22] and [24] have the longest running time, which is due to the fact that they both improve on the traditional cosine similarity algorithm. The recommendation efficiency of [24] is less than that of [22] because the algorithm of [24] takes into account the changes in student interest level that occur with time offset. The recommendation efficiency of the algorithm in this paper is close to that of [21] and [23], and [21] dominates relative to the dataset selected in this paper. However, since the fusion algorithm incorporates [24] algorithm, it is longer than the algorithms of [23] and [21] in terms of running time. Overall, the proposed algorithm can meet the basic needs of students.

## 4. Conclusion

Since China's reform and opening up, its communication with foreign countries has become more and more frequent. As the most widely used language in the world, English plays a crucial role in China's communication with other countries. Therefore, English teaching and learning are very important. Writing ability is one of the most important and difficult to develop. The main way to improve college students' English writing ability is to read a large amount of excellent English writing literature. However, based on the varying level of teachers, reading amount, and experience, it is impossible for reference recommendations to meet the writing needs of all students. How to meet the needs of students with different English levels, provide accurate and personalized reference books for each college student in a mass of references, and make real-time recommendation through online recommendation system is the top priority in universities. Therefore, this paper proposes a hybrid recommendation algorithm based on multidimensional feature representation learning (MFL). The algorithm split the English writing literature scoring network and, based on the improved LINE algorithm, carried out hierarchical advance learning for college students who like English writing literature and dislike English writing literature. Based on the improved DeepWalk algorithm, the similar student sequence is obtained and the similar student features are captured. After linear combination of liking features, dislike features, and similar students' features, the final feature vector of students was connected, and cosine similarity was used as the similarity measurement index to achieve the recommendation of English writing literature. Experimental results show that the algorithm can not only take into account the multidimensional nature of students and English writing literature, but also improve the efficiency and effectiveness of recommendation. Nowadays, recommendation systems have been integrated with all aspects of life, work, and study, and the next research goal is to apply them to other disciplines or to trigger more comprehensive recommendations from the intersection of multiple disciplines.

## Data Availability

The labeled dataset used to support the findings of this study can be obtained from the corresponding author upon request.

## Conflicts of Interest

The authors declare no conflicts of interest.

## Acknowledgments

This study was supported by Zhengzhou Business University.

## References

[1] M. Cochran-Smith, L. Grudnoff, L. Orland-Barak, and K. Smith, "Educating teacher educators: international perspectives," *The New Educator*, vol. 16, no. 1, pp. 5–24, 2020.

[2] I. Suastra and S. Menggo, "Empowering students' writing skill through performance assessment," *International Journal of Language Education*, vol. 4, no. 3, pp. 432–441, 2020.

[3] Z. Turan and B. Akdag-Cimen, "Flipped classroom in English language teaching: a systematic review," *Computer Assisted Language Learning*, vol. 33, no. 5-6, pp. 590–606, 2020.

[4] W. S. Albiladi and K. K. Alshareef, "Blended learning in English teaching and learning: a review of the current literature," *Journal of Language Teaching and Research*, vol. 10, no. 2, pp. 232–238, 2019.

[5] M. A. Ghufron and S. Ermawati, "The strengths and weaknesses of cooperative learning and problem-based learning in EFL writing class: teachers and students' perspectives," *International Journal of Instruction*, vol. 11, no. 4, pp. 657–672, 2018.

[6] T. M. Al-Jarrah, J. M. Al-Jarrah, R. H. Talafhah, and I. Bashir, "Exploring the effect of metacognitive strategies on writing performance," *Global Journal of Foreign Language Teaching*, vol. 9, no. 1, pp. 33–50, 2019.

[7] C. J. Lin, G. J. Hwang, and Q. K. Fu, "A flipped contextual game-based learning approach to enhancing EFL students' English business writing performance and reflective behaviors," *Journal of Educational Technology & Society*, vol. 21, no. 3, pp. 117–131, 2018.

[8] X. Wang, "Literature review on language learning strategy research in China," *Journal of Language Teaching and Research*, vol. 11, no. 5, pp. 787–792, 2020.

[9] C. Zhou, "Alleviation of psychological anxiety of college English learners in writing through blended teaching," *Revista Argentina de Clinica Psicologica*, vol. 29, no. 2, p. 73, 2020.

[10] W. Chen and S. Yu, "Implementing collaborative writing in teacher-centered classroom contexts: student beliefs and perceptions," *Language Awareness*, vol. 28, no. 4, pp. 247–267, 2019.

[11] P. J. B. Tan, "An empirical study of how the learning attitudes of college students toward English E-tutoring websites affect site sustainability," *Sustainability*, vol. 11, no. 6, p. 1748, 2019.

[12] L. Noorizadeh-Honami and A. Chalak, "Comparative analysis of architecture research article abstracts written by native and non-native authors: a cross-linguistic, cross-cultural study," *Theory and Practice in Language Studies*, vol. 8, no. 3, pp. 325–330, 2018.

[13] S. Graham, K. Wijekumar, K. R. Harris et al., "Writing skills, knowledge, motivation, and strategic behavior predict students' persuasive writing performance in the context of robust writing instruction," *The Elementary School Journal*, vol. 119, no. 3, pp. 487–510, 2019.

[14] L. S. Teng, P. P. Sun, and L. Xu, "Conceptualizing writing self-efficacy in English as a foreign language contexts: scale validation through structural equation modeling," *Tesol Quarterly*, vol. 52, no. 4, pp. 911–942, 2018.

[15] N. Us Saqlain, A. Shafqat, and A. Hassan, "Perception analysis of English language teachers about use of contextualized text for teaching ESP," *Asian ESP Journal*, vol. 16, no. 5.1, pp. 275–299, 2020.

[16] S. Canagarajah, "The unit and focus of analysis in lingua franca English interactions: in search of a method," *International Journal of Bilingual Education and Bilingualism*, vol. 21, no. 7, pp. 805–824, 2018.

[17] S. Nazir, F. Abbas, and F. Naz, "Historical development of orthography in English and impact of computer-mediated communication (CMC) on the emerging orthographic patterns in English," *PalArch's Journal of Archaeology of Egypt/Egyptology*, vol. 17, no. 11, pp. 162–175, 2020.

[18] E. Rafatbakhsh and A. Ahmadi, "A thematic corpus-based study of idioms in the Corpus of Contemporary American English," *Asian-Pacific Journal of Second and Foreign Language Education*, vol. 4, no. 1, pp. 11–21, 2019.

[19] P. S. Rao, "The importance of speaking skills in English classrooms," *Alford Council of International English & Literature Journal (ACIELJ)*, vol. 2, no. 2, pp. 6–18, 2019.

[20] J. Li, "A corpus-based study on content validity of proofreading tasks in gaokao," *Multimedia-Assisted Language Learning*, vol. 25, no. 1, pp. 31–55, 2022.

[21] I. I. Kostikova and S. V. Miasoiedova, "Supporting postgraduate students writing skills development with the online

machine learning tool: write & improve," *Information Technologies and Learning Tools*, vol. 74, no. 6, pp. 238–249, 2019.
[22] A. Farouk and D. Zhen, "Big data analysis techniques for intelligent systems," *Journal of Intelligent and Fuzzy Systems*, vol. 37, no. 3, pp. 3067–3071, 2019.
[23] G. M. A. Grami, "An evaluation of online and automated English writing assistants: collocations and idioms checkers," *International Journal of Emerging Technologies in Learning (IJET)*, vol. 15, no. 4, pp. 218–226, 2020.
[24] X. Chen, D. Zou, and H. Xie, "Two decades of artificial intelligence in education," *Educational Technology & Society*, vol. 25, no. 1, pp. 28–47, 2022.