International Journal for Housing Science and Its Applications
Publish August 6, 2025. Volume 46, Issue 3 Pages 5103-5117

https://doi.org/10.70517/ijhsa463438

Research on the Automated Assistance Platform for English Business Writing Based on Grammar Analysis Algorithm

Zhenying Zhang[1,*]
[1] Basic Teaching Department, Shangqiu Institute of Technology, Shangqiu, Henan, 476000, China
Corresponding authors: (e-mail: 1350013064@sqgxy.edu.cn).

Abstract With the acceleration of globalization, the importance of business English writing in cross-cultural communication is becoming more and more prominent. This study proposes an automated assistance platform for English business writing based on grammatical analysis algorithms, which deeply integrates the RST-Style discourse parser improved by the Conditional Random Field CRF with the GloVe global semantic word vector model to solve the deficiencies of traditional methods in long-distance dependency and lexical semantic association, and introduces a sequence-to-sequence error correction model based on the replication mechanism combined with the BERT pre-training language model to optimize the semantic representation and error correction efficiency. Through multi-dimensional experimental validation, the model has an average absolute error MAE of 2.071 and a Pearson's correlation coefficient PCC of 0.702 in the lexical articulation diagnosis task.The pairwise accuracy PRA for logical coherence diagnosis on the Accident and Earthquake datasets are 96.57% and 97.98%, respectively. The F1 value for the grammatical error detection task reaches 69.84%, which is significantly better than the baseline model. The teaching application experiments show that the mean of the total posttest score of the experimental group using the platform improves to 90.41 (58.87 on the pre-test), and the subdimensions of lexical articulation and grammatical accuracy are close to full scores of 23.16 and 24.01, respectively, and the standard deviation is significantly narrowed, which confirms the practical value of the platform in improving writing ability and teaching efficiency.

Index Terms grammatical analysis, English business writing, automated assistance platform, conditional random field, RST-Style

I. Introduction

English writing is an important part of language learning, which not only requires students to use vocabulary accurately, but also to master and flexibly utilize appropriate grammatical structures to express ideas and convey information [1]. The purpose of teaching business English writing is to improve the fluency and appropriateness of students' writing expressions in various business English situations [2]. By creating real business and management scenarios and case studies, students' knowledge and proficiency in the language of business English can be enhanced and their professional competitiveness in the international business environment can be strengthened [3], [4]. The teaching content includes writing training in scenarios such as workplace, business traveling, team building, trade negotiation and finance and economics [5], [6].

At present, in the teaching of business English writing, the teaching of grammatical structures often focuses on the memorization of rules and mechanical exercises, neglecting their application and optimization in specific contexts [7]. This teaching method easily leads to the problem of “grammatically correct but hard expression” in actual writing, which affects the fluency and readability of the article [8], [9]. However, technological advances are helping the various specialized fields of higher education to continuously extend and expand, so that learners can continuously acquire new knowledge and skills beyond the books [10]. The establishment of an automated writing assistance platform can help students understand sentence structure at a theoretical level through detailed grammar explanations combined with vivid example sentences demonstrating the correct usage of sentence patterns and contexts of use [11]-[13]. Meanwhile, in supplemental instruction, students can get writing materials provided by the platform as well as opportunities to interact with other members of the same group, and these interactions in turn develop their communication skills and collaboration skills [14]-[16]. Therefore, exploring appropriate writing analysis algorithms and automated assistance platforms is an important way to promote the quality improvement of English business writing education in colleges and universities.

This study focuses on the core technology model for building an automated assistance platform for business English writing, aiming to solve the key problems of grammatical standardization, semantic coherence and error correction efficiency in business English writing through the deep integration of grammar analysis algorithms and deep learning models. The article firstly starts from the characteristics of business English writing and the advantages of network-assisted teaching, and emphasizes the enhancement of learner's initiative and teaching flexibility under the empowerment of technology. On this basis, two core grammar analysis models are proposed: the RST-Style discourse parser based on the improvement of Conditional Random Field CRF and the global semantic-driven GloVe word vector model. The former optimizes syntactic structure parsing by capturing contextual dependencies, solves the traditional model's neglect of long-distance dependencies, and optimizes tag prediction through global features. The latter enhances lexical semantic representation from the perspective of co-occurrence probability, quantifies the strength of semantic associations between words by counting the frequency of word co-occurrence in the corpus. And a weighting function is introduced to balance the influence of high-frequency words and low-frequency words. Matrix decomposition and training are also carried out to generate word vectors with global semantics by weighted least squares decomposition of the co-occurrence matrix. Finally, to address the contradiction between local error correction needs and generation flexibility in business writing, a sequence-to-sequence error correction model based on the replication mechanism is introduced, which is combined with the initialization and fusion strategy of the BERT pre-trained language model. The decoding dynamically chooses to copy the vocabulary from the source text or generate new words to avoid excessive modification of the original semantics. And the replication probability is calculated by attention weights to ensure that the error correction focuses on the error fragments.BERT-init initializes the model with pre-trained BERT parameters and preserves the generic semantic knowledge.BERT-fuse takes the BERT encoding results as additional feature inputs to enhance the contextual characterization capability and achieve efficient and accurate syntactic error correction.

II. Research on model construction and grammar analysis algorithm of business English writing automation assistance platform

II. A. Characteristics of Business English Writing and Advantages of Web-Assisted Writing Teaching

The teaching of modern business English writing, in order to prevent the teaching method from flowing on the surface of the language, one-sidedly focusing on the expression of grammar and vocabulary use, but also to take into account the inherent mode of expression of business etiquette and the improvement of the students' practical ability, puts forward higher requirements for both teachers and students. The popularization of technology provides students with opportunities for active learning. Students collect writing materials on the Internet, can obtain the maximum information capacity within a unit of time, change passive learning into active learning, and realize their own main position in the teaching process. Teachers through e-mail and other ways to present the content of the composition, teaching writing methods, organizing online discussions, individual tutoring of students, and is not subject to time and classroom and other conditions, greatly improving the flexibility and diversity of teaching. Relevant scientific research has proved that 11% of people's learning is carried out through hearing and 83% through vision. That is to say, the “attention input” information acquired by people through vision and hearing is 94% of all the information acquired. From the perspective of memory, the use of both visual and auditory senses in the learning process can significantly improve the learning efficiency and memory effect. The use of network technology is to maximize students' learning of English, to accelerate the formation of English thinking and shorten the process of “mental translation”. At the same time, the use of the network for English teaching can be random and multi-dimensional information to vividly present language knowledge to the students, thus mobilizing their multi-sensory system, so that they actively participate in the teaching of business English writing, breaking through the difficulties and key points in English teaching, optimizing the teaching process.

II. B. English Grammar Analysis Model

Although web-assisted teaching significantly improves the flexibility and interactivity of business English writing, its effectiveness still relies on the underlying technology to accurately analyze the grammatical structure and semantic expression. To this end, this section will explore in depth the core grammar analysis model of the supporting platform in order to solve the pain point of insufficient grammatical standardization in traditional teaching.

II. B. 1) RST-Style Discourse Parser Based on CRF Improvement

On parsing methods for machine learning and deep learning. In this paper, we will take the classic RST-Style discourse parser framework and improve on it to parse text using the Conditional Random Field (CRF) method.

A random field is a mathematical model used to model relationships between multiple random variables, usually representing a graph structure consisting of multiple locations or nodes. Each location or node in a random field corresponds to one or more random variables, and each random variable has a certain probability of taking on a certain value; the entire set is called a random field. Markov random field is a special kind of probabilistic graph model, its basic assumption is that the value of each node is only related to its direct neighbors and will not be affected by other nodes, this assumption simplifies the process of modeling random fields.The conditional assumption of CRF probabilistic undirected graph model is based on the Markov random field composed of random variables, and unlike the premise assumption, it does not pay attention to the union between the output variables distribution and focuses mainly on the conditional dependence between the input and output variables. The mathematical language of CRF is described as assuming that X is a random variable in the observation sequence and Y is a random variable in the output sequence, and that the probability distribution of Y given X and Y is $P(Y|X)$, and that the conditional probability distribution of $P(Y|X)$ is a conditional random field when the random variable Y constitutes a Markov random field. The structure of X and Y in CRF is not necessarily the same, when X and Y have the same structure i.e. each node corresponds to the same location, the CRF is converted into a linear chain conditional random field.

The linear chain conditional random field is divided into two types of eigenfunctions, the state eigenfunction is defined on the Y node, as shown in Equation (1), which is only related to the state of the current node and does not need to take into account the contextual information, i is the current node's position in the sequence, and L is the total number of eigenfunctions defined on the node. Another type of feature function is the transfer feature function defined in the context of node Y as shown in equation (2), this type of feature function is only related to the current position and the position of the previous node, K is the total number of transfer features defined at that node.

$$s_l(y_i, x_i, i), i = 1,2,...L \tag{1}$$

$$t_k(y_{i-1}, y_i, x, i), k = 1,2,...K \tag{2}$$

Linear chain conditional random field is a model based on Markov property, which only focuses on the relationship between neighboring nodes and does not consider the relationship between non-neighboring nodes. Regardless of the state eigenfunctions or transfer eigenfunctions, they take non-zero or one values, indicating whether the current node and neighboring nodes satisfy the eigenconditions. Each feature function corresponds to a weight value, and the model automatically learns the weight value of each feature to more accurately predict the label or attribute of unknown data. The parameterized form of the linear chain conditional random field is shown in Equation (3).

$$P(y|x) = \frac{1}{Z(x)} \exp\left(\sum_{i,k} \lambda_k t_k(y_{i-1}, y_i, x, i) + \sum_{i,l} u_l s_l(y_i, x_i, i)\right) \tag{3}$$

$$Z(x) = \sum_y \exp\left(\sum_{i,k} \lambda_k t_k(y_{i-1}, y_i, x, i) + \sum_{i,l} u_l s_l(y_i, x_i, i)\right) \tag{4}$$

λ_k, u_l are two kinds of eigenfunction weights, respectively, and $Z(x)$ is the normalization factor, as shown in Equation (4). The conditional probability distribution of a linear chain conditional random field can be viewed as a combination consisting of the rules of all the eigenfunctions and their corresponding weights, which describe the relationship between individual states in the sequential data and the relationship between states and observations. By using these rules and weights, a linear chain conditional random field can infer the most likely sequence given the observations. In a linear chain conditional random field, the probability of each node depends on the input sequence X as well as the outputs $Y-1$ and $Y+1$ of neighboring nodes, a modeling approach that takes into account the linguistic regularity and contextual semantics of each vocabulary. Conditional random fields are a discriminative model-based modeling approach that can take transfer probabilities between long-distance contexts into account in the labeled sequences and use more diversified global features for global parameter optimization and decoding, an operation that solves the problem of label bias that exists in other discriminative models.

II. B. 2) The word vector model GloVe

Although CRF-based discourse parsers can effectively model contextual grammar rules, their capture of lexical semantics still needs to rely on finer-grained representations. For this reason, this section further introduces the GloVe word vector model, which complements the deeper associations of lexical semantics through the decomposition of the global co-occurrence probability matrix, forming a dual parsing framework of syntax and semantics.

GloVe is a pre-trained model of words represented as vectors, compared to the Word2Vec model above which utilizes local contextual features for modeling, GloVe aims to obtain semantic features between words from a global perspective and apply them to downstream tasks.GloVe represents the semantic information of words by transforming and decomposing vectors into a co-occurrence probability matrix into a word vector matrix, and uses matrix decomposition techniques to obtain the final word vector representation. If two words frequently occur in the same context, the higher their co-occurrence probability is, the closer their semantic relationship is. An overview of the principles of GloVe is given next.

First, each word in the corpus is scanned to build a co-occurrence matrix M, and each word corresponds to a vector M_i, which represents the total number of times word i occurs in the context in the corpus. The co-occurrence matrix M_{ij} represents the ratio of the number of times word i appears around word j to the number of times word i appears in the whole corpus, the larger the ratio is the closer the relationship between word i and word j, and the formula is shown in (5). P_{ij} denotes the co-occurrence probability of word i appearing in the environment of word j, then P_{ij} is computed as shown in equation (6) below:

$$M_i = \sum_{j=1}^{N} M_{ij} \tag{5}$$

$$P_{ij} = P(j \mid i) = \frac{M_{ij}}{M_i} \tag{6}$$

Assume that the word vectors of word i, word j and word k are denoted as v_i, v_j and v_k, respectively, and denote the probability of word k appearing in the context of word i by P_{ik}, and the probability of word k co-occurring in the context of word j by P_{jk}.The core idea of the GloVe model is to infer semantic correlations among words based on the probability of their co-occurrence in the corpus. The core idea of GloVe model is to infer the semantic correlation between words based on their co-occurrence probability in the corpus. In order to make the word vectors contain the information of the co-occurrence matrix, we use the probability scaling function F to compare whether word k is more relevant to word i or word j. The exact formula is shown in equation (7) below:

$$F(v_i, v_j, v_k) = \frac{P_{ik}}{P_{jk}} \tag{7}$$

The similarity of two vectors is evaluated in a linear space, and the parameters in the F function are represented as vector inner products after doing difference operations, and this transformation when training the model can highlight more the proportional difference in the co-occurrence probability between the words, thus representing the semantic relationship between the words. Equation (8) is obtained as follows:

$$F((v_i - v_j)^T v_k) = F(v_i^T v_k - v_j^T v_k) = \frac{P_{ik}}{P_{jk}} \tag{8}$$

In order to balance the weights of high-frequency and low-frequency words in word vectors and better reflect the importance of words in the corpus, a weight function f is introduced, as shown in the following equation (9), and according to the experience when $M_{\max}$ takes the value of 100, in order to make the weights more reasonable, the value of α should be 3/4. The construction of the weighted least squares cost function is obtained as shown in Eq. (10), with the aim of avoiding excessive weighting of high-frequency words so that they can be adjusted for word frequency.

$$f(M_{ij}) = \begin{cases} \left(\dfrac{M_{ij}}{M_{\max}} \right)^{\alpha} & M_{ij} < M_{\max} \\ 1 & otherwise \end{cases} \tag{9}$$

$$J = \sum_{i,j=1}^{N} f\left(M_{ij}\right)\left(v_i^T v_j + b_i + b_j - \log\left(M_{ij}\right)\right)^2 \tag{10}$$

GloVe uses stochastic gradient descent to update the model parameters and minimizes the cost function to obtain an optimal word vector representation for subsequent downstream tasks. This method can better capture the semantic relationships between words and adapt to different textual contexts than methods that use only local contexts or only global word co-occurrence matrices.

II. C.Sequence-to-sequence error correction model

The improvement of syntactic analysis and semantic representation provides the foundation for writing assistance, however, the practical application still requires dynamic correction of local errors in the text. Based on this, this section proposes a sequence-to-sequence error correction model, which combines the replication mechanism with the BERT pre-trained language model to achieve efficient error correction while maintaining the intent of the original text, and ultimately forms a closed-loop technology link from parsing to correction.

Although the NMT method has reached the optimal level (SOTA) in the field of GEC, the error correction task is different from the translation task in that it mainly modifies a few words in the source sentence, and most of the words are kept unchanged, in view of which the grammatical error correction requires a more appropriate neural network architecture. In this paper, we propose a GEC model based on the replication mechanism, whose core idea is to consider two different generative distributions in the process of generating sequences: the first is the probability distribution of replicating the words in the input sequences, and the second is the probability distribution of generating the words from the candidate lexicon. In short, the copying mechanism decides whether to copy words directly from the source text or not, while the other mechanism selects new words from the lexicon to generate. These two probability distributions are combined to form the final generation probability distribution. This is done by weighting and summing the two probability distributions and predicting the words that should be generated at each moment based on this combined probability distribution. The replication mechanism's can be divided into several key components:

Probabilistic mixture model: the final output probability distribution P_t is a mixture of the generation distribution P_{gen_t} and the copying distribution P_{copy_t}. This model allows the system to flexibly switch between generating new words and copying existing words in the source sentence. The formula is shown in equation (11):

$$p_t(w) = (1-\alpha_{copy_t}) \times p_{gen_t}(w) + (\alpha_{copy_t}) \times p_{copy_t}(w) \tag{11}$$

where α_{copy_t} is the factor used to balance copying and generation, and its value range is $[0,1]$.

Attention distribution: the calculation of the copying score relies on the new attention distribution between the current hidden state of the decoder and the hidden state of the encoder. Specifically, this attention distribution is computed as follows:

$$q_t, K, V = h_{trg_t} W_q^T, H_{src} W_k^T, H_{src} W_v^T \tag{12}$$

$$A_t = q_t^T K \tag{13}$$

$$P_{copy_t}(w) = softmax(A_t) \tag{14}$$

where q_t, K and V represent the query, key and value, respectively, which are the elements needed to compute the attention distribution and replicate the hidden state. The balancing factor is computed as follows.

$$\alpha_{copy_t} = sigmoid\left(W^T X\left(A_t \cdot V\right)\right) \tag{15}$$

This factor determines whether, at a given time step t, the system tends to copy words from the source sentence or to generate words from the glossary.

An intriguing question in the task of syntactic error correction is whether sequence-to-sequence based GEC models can benefit from recent advances in masked language modeling (MLM). Common strategies for incorporating MLM into EncDec models include initialization (init) and fusion (fuse). In the initialization approach, the downstream task model starts with the parameters of the pre-trained MLM and is subsequently trained on a task-specific training set. However, this approach has limited effectiveness in tasks such as sequence-to-sequence language generation, which typically require large amounts of task-specific training data, and fine-tuning the MLM on such a large dataset often results in the corruption of its pre-trained representations, which triggers catastrophic oblivion. On the other hand, in the fusion approach, the pre-trained representation of the MLM is used as an additional feature when training the task-specific model. When applying this approach to GEC, what the MLM learns during pre-training is preserved; however, the MLM does not adapt to the GEC task or to a specific input distribution, which may limit the potential of GEC models to effectively utilize the MLM.

In this paper, we investigate how to effectively incorporate BERT into GEC tasks, which is categorized into the following two main approaches.

BERT-init: directly using the pre-trained model parameters to initialize the corresponding parameters in the new model, followed by fine-tuning.

BERT-fuse: it takes an input sentence of length n $X=(x_1,...,x_n)$, where x_i denotes the i th token in X. First, BERT encodes the input sentence and produces the representation $B=(b_1,...,b_n)$. Subsequently, the GEC model encodes X and B as inputs. In the GEC model encoder, the i th hidden representation of the l th layer is denoted as $h_{l_i} \in H$, where h_0 denotes the word embedding of the input sentence X. Next, the formula for computing h_{l_i} is adjusted as follows:

$$\tilde{h}_{l_i} = \frac{1}{2}\left(\mathrm{A}_h\left(h_{l-1}^i, \mathrm{H}_{l-1}\right) + \mathrm{A}_b\left(h_{l-1}^i, \mathrm{B}_{l-1}\right)\right) \tag{16}$$

Here, A_h and A_b are the hidden layer H of the GEC encoder and the attention model of the BERT output B, respectively. Each $\tilde{h}_{l_i}$ is further processed by the feed-forward network F to produce the l th layer $\mathrm{H}_l = (\mathrm{F}\ (\tilde{h}_l^1),...,\mathrm{F}\ (\tilde{h}_l^n))$. The hidden state $s_{l_t} \in S$ of the decoder is computed as:

$$\hat{s}_l^t = \mathrm{A}_s\left(s_{l-1}^t, S_{l-1}^{<t+1}\right) \tag{17}$$

$$\tilde{s}_l^i = \frac{1}{2}\left(\mathrm{A}_h(\hat{s}_{l-1}^i, \mathrm{H}_{l-1}) + \mathrm{A}_b(\hat{s}_{l-1}^i, \mathrm{B}_{l-1})\right) \tag{18}$$

$$s_l^t = \mathrm{F}\ (\tilde{s}_l^t) \tag{19}$$

where A_s represents the self-attention model. Eventually, s_{L_t} is processed by linear transformation and softmax function to predict the t th word $\hat{y}_t$.

III. Performance testing of grammatical analysis models and coherence diagnostic experiments

Chapter 2 constructs a grammar analysis model based on CRF and GloVe, and proposes a sequence-to-sequence error correction framework. In order to verify the practical efficacy of the above models in business English writing, Chapter 3 will systematically evaluate their performance in lexical articulation, logical coherence and grammatical error correction tasks through multi-dimensional experiments.

III. A. Experimental setup

III. A. 1) Experimental environment

Hardware Environment Processor: Intel(R) Core(TM) i5-3470 CPU @3.20GHz 3.60 GHz. 16GB of RAM. The software environment operating system is Microsoft Windows 10 64-bit. Development languages used are Java and Python. The development environment is Java JDK1.8, Python3.6. Development tools are Eclipse, PyCharm.

In this paper, CRF-based improved RST-Style employs GloVe to generate 300-dimensional word vectors for each word of the clause as the initial features of the word, and embedding representation of the clause by vector averaging. After constructing the required feature matrix and adjacency matrix, two layers of stacked GCN are used to process the matrix, in order to improve the generalization ability of the model, a layer of Dropout is attached to each layer of GCN and set to 0.5, the experimental optimizer is Adam, the weight decay coefficient is 5e-5, and the number of Epochs is set to 50.Finally, the softmax function is used to predict the English text Logical coherent classification and the corresponding probability are obtained.

III. A. 2) Experimental data sets

In the writing lexical articulation diagnostic experiment, the CRF-improved RST-Style discourse parser incorporating GloVe obtained lexical articulation scores for English texts by fusing entity distribution scores and articulation scores.In order to evaluate the performance of this paper's model on the task of scoring for discourse coherence, the Corpus of English Language Learners (CELC) was used as the model's dataset for the purpose of this chapter.The CELC corpus was extracted from the writing output of university English majors, university non-English majors, and secondary English learners, and each English text in this corpus was evaluated by several professional English teachers with a final overall rating of the English text. As one of the important factors affecting the scoring of English texts, the scoring of the coherence quality of English texts can also be reflected from the teachers' overall scores of the English texts. The English texts in the CELC corpus are scored out of 100, and when the higher the score is, it means that the coherence of the English text is relatively higher.

In the writing vocabulary articulation diagnostic experiment, the core of the computational entity distribution score of the English grammar analysis model designed in this paper is to capture and analyze the distributional differences of grammatical role transfer sequences in coherent and incoherent English texts, specifically, the model is firstly trained by coherent texts to learn and generate distributional features of grammatical role transfer sequences with high coherence quality, and then features matching is carried out for the English texts to be diagnosed in the test set, and finally, features matching is carried out to obtain the distributional features of grammatical role transfer sequences of English texts in the test set. English texts, then match the features to the English texts in the test set, and finally obtain the entity distribution scores of the English texts. Based on this, this chapter selects 1,000 English texts with high overall scores from the CELC corpus as the training set for the writing vocabulary articulation diagnostic experiments, and then another 300 English texts with score differences in this corpus are selected as the test set for the writing vocabulary articulation diagnostic experiments, to validate the validity and accuracy of the model in evaluating the coherence of English texts.

In the logical coherence diagnostic experiment, the English grammatical analysis model designed in this paper obtained the coherence probability distributions of English texts by parsing the logical relationships within the texts.In order to evaluate the performance of the model in the discourse coherence identification task, two public datasets were used in this chapter: Accident and Earthquake.The Accident dataset consists of reports of airplane crashes from the U.S. National Transportation Safety Board, and the Earthquake dataset consists of reports of airplane crashes from the U.S. National Transportation Safety Board. The Accident dataset consists of reports about airplane crashes from the National Transportation Safety Board, with an average length of 11.7 sentences, and the whole dataset is divided into three parts: the training set, the validation set, and the test set, with 100 articles in each of the training and test sets, of which 10 articles in the training set are randomly taken as the initial validation set; the Earthquake dataset consists of news about earthquakes from the Associated Press, with an average length of 10.2 sentences, and the whole dataset is divided into the training set, validation set, and test set. The dataset is divided into three parts: training set, validation set and test set, with a total of 100 articles in the training set and 99 articles in the test set, of which 10 articles are randomly taken in the training set as the initial validation set. Then all the articles in the training set, validation set, and test set of the two datasets are randomly disrupted in the order of statements 20 times respectively to obtain the comparison articles that do not repeat each other (some articles are shorter resulting in the generation of less than 20 articles). In the end, the Accident dataset generated a total of 4,274 pairs of original-text-disordered articles, totaling 4,425 articles; the Earthquake dataset generated a total of 4,037 pairs of original-text-disordered articles, totaling 4,329 articles.

The ratio of training, validation, and test sets for both datasets is close to 10:1:10, where the test set for the Accident dataset generated a total of 2011 pairs of original-text-disordered texts, and the test set for the Earthquake dataset generated a total of 1,974 pairs of original-text-disordered texts.

III. A. 3) Assessment of indicators

In the writing lexical articulation diagnostic experiment, the model diagnoses the lexical articulation scores of English texts. In order to assess the reasonableness of the model's scoring, two metrics, Mean Absolute Error (MAE) and Pearson's Coefficient (PCC), are used in this chapter to carry out the evaluation of the experimental results.

The mean absolute error is an index used to quantify the degree of deviation between the predicted value and the actual value, which is calculated as the average of the absolute value of the difference between the model's predicted score X and the actual score Y of the English text, and N is the total number of the English text. The mean absolute error does not involve the mutual offsetting of the errors, so it can reflect the size of the prediction error more realistically and accurately, and in general, the smaller the mean absolute error is, the the more accurate the prediction model is. The formula is shown in (20).

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|X_i - Y_i\right| \tag{20}$$

Pearson's coefficient is a statistical index used to measure the strength and direction of the linear relationship between two variables, put in this experiment, Pearson's coefficient is obtained by calculating the quotient of the covariance and standard deviation of the two variables of the model's predicted rating X, the actual rating of the English text Y, in which X and Y are the means of the two variables, respectively, and the Pearson's coefficient ranges from -1 to 1, in which the value of 1 indicates that the two variables are completely positive correlation, while -1 indicates complete negative correlation, and 0 indicates no correlation. In the experiments in this chapter, the closer the Pearson coefficient is to 1, the more accurate the prediction model is. The calculation formula is shown in (21).

$$PCC = \frac{\sum_{i=1}^{N}\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\sqrt{\sum_{i=1}^{N}\left(X_i - \overline{X}\right)^2}\sqrt{\sum_{i=1}^{N}\left(Y_i - \overline{Y}\right)^2}} \tag{21}$$

In the Writing Logic Coherence Diagnostic Experiment, the model diagnoses the coherence probability distribution of English texts, following the evaluation criteria of the Accident dataset and the Earthquake dataset, this chapter adopts the Pairwise Accuracy Rate (PRA) as an evaluation metric for this experiment, which is placed in the above mentioned dataset, and represents the proportion of the N pairs of original text-gone-wrong texts of the test set that can be correctly recognized as the The percentage of data pair NUMaccuracy of the original text, comparing the coherence probability of a pair of original text and disordered text, if the coherence probability of the original text is higher than that of the disordered text, it means that this data pair is accurately recognized as the original text. The specific formula for pairwise accuracy is shown in (22).

$$PRA = \frac{NUM_{accuracy}}{N} \tag{22}$$

III. A. 4) Comparison of baseline models

In order to test the performance of the models proposed in this paper on the tasks of writing vocabulary diagnosis and writing logic diagnosis identification, a series of classical and representative syntactic analysis models are selected for comparative experiments on the CELC dataset, Accident and Earthquake datasets.

Graph: this model is a graph-based local coherence diagnostic party, which realizes the discourse coherence diagnosis of text by representing the entity transformation information in the form of a graph, and improves the computational efficiency of the model.

CSM: This model is another coherence diagnostic model proposed on the basis of entity grid model.

HMM: This model is a coherence diagnosis model based on syntactic structure, which learns syntactic structure features and then diagnoses the coherence of English texts.

Recursive: a neural network architecture for coherence tasks based on distributed sentence representations, of which the recurrent neural network model is one, which automatically learns the syntactic and semantic representations of sentences to diagnose the coherence of texts.

Recurrent: a recurrent neural network model that can also diagnose text coherence by automatically learning syntactic and semantic representations of sentences.

s-CDI: a fast coherence detection (FFCD) model based on BERT NSP2, in which the short text coherence deficit index (s-CDI) is defined for assessing the coherence of short texts.

SEDG: a discourse coherence analysis model combining sentence embedding and dimensional grids, which utilizes deep learning techniques to obtain sentence-level vector representations, and cleverly combines sentence embedding with dimensional grids to effectively model the discourse coherence of texts.

III. B. Comparative experiment on writing vocabulary articulation diagnosis

Based on the above experimental environment and dataset configuration, this section compares the performance of different models in the lexical articulation diagnostic task, focusing on analyzing the MAE and PCC metrics to validate the enhancement of semantic associations by GloVe word vectors. The comparison of the experimental results of different models on the two assessment metrics of MAE and PCC on the CELC dataset is shown in Table 1.

Table 1: The performance of different models in discourse coherence scoring

Model	MAE	PCC
Graph	4.947	0.473
CSM	4.357	0.454
HMM	4.124	0.429
Recursive	3.427	0.592
Recurrent	3.455	0.607
s-CDI	2.981	0.599
SEDG	2.241	0.633
OURS	2.071	0.702

In the lexical articulation diagnosis task, the CRF-RST-Style model proposed in this paper significantly outperforms other comparison models on the CELC dataset. Its mean absolute error MAE is only 2.071, which is much lower than that of the traditional feature engineering model Graph model (4.947) and the deep learning model s-CDI model (2.981.) At the same time, the Pearson correlation coefficient (PCC) of this paper's model reaches 0.702, which indicates that its predicted scores are highly positively correlated with the manual scores. In contrast, although the RNN-based model Recurrent performs better in PCC at 0.607, its mean absolute error MAE is 3.455, which is still significantly larger than that of this paper's model, suggesting that this paper's model is more advantageous in taking into account the error control and correlation.

III. C. Writing Logical Coherence Diagnostic Comparison Experiment

The lexical articulation experiment verifies the model's ability to analyze local semantics, and in order to further evaluate its effectiveness in capturing global logical relations, this section tests the model's pairwise accuracy PRA in the Accident and Earthquake datasets, revealing the advantages of the improved CRF algorithm in long-distance dependency parsing.

In this experiment, PRA is used as an evaluation metric to continue the comparison experiment with the above models. The above models can be divided into two main categories, the first three models belong to feature engineering-based models, and the last four models belong to deep learning-based models. The specific experimental results about logical coherence diagnosis of discourse coherence are shown in Figure 1, and the data in the table are presented in the form of percentage %.

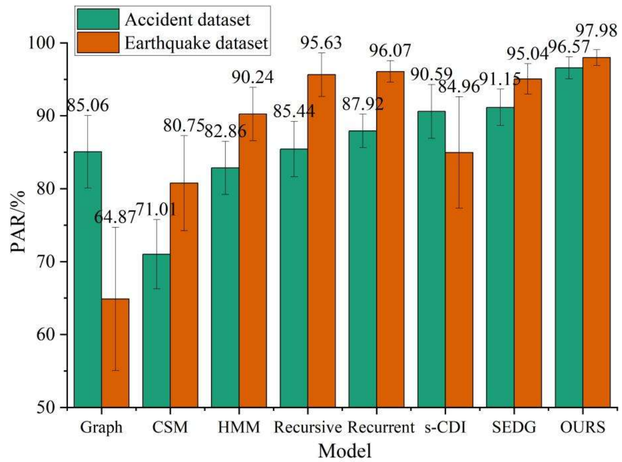

Figure 1: Experimental results on logical coherence in diagnosing discourse coherence

In the logically coherent diagnosis task, the pairwise accuracy PRA of this paper's model on the two datasets Accident and Earthquake reaches 96.57% and 97.98%, respectively, which are both ahead of all baseline models. Especially on the Accident dataset, this paper's model improves by 5.42 percentage points over the 91.15% of the second-place SEDG model, showing stronger robustness. Notably, the HMM model outperforms the Earthquake dataset with 90.24%, but it is still lower than the 97.98% of this paper's model, indicating that the combination of global semantic-driven GloVe word vectors and CRFs effectively captures long-distance logical dependencies.

III. D. Comparison experiment between model and teacher manual scoring

The logical coherence experiment proves the automated scoring ability of the model, and in order to verify its consistency with the actual teaching requirements, this section will compare the model's predictive scoring with the teacher's manual scoring, quantify the correlation between the two, and provide empirical evidence for the landing of the platform.

In order to evaluate the application effect of the CRF-RST-Style model incorporating GloVe word vectors proposed in this paper, the research group invites professional English teachers to formulate a standardized manual scoring criterion for English text coherence according to the requirements of English text coherence scoring in terms of vocabulary articulation, logical coherence, etc., and to conduct scoring of 1000 English texts in the CELC Corpus in accordance with the criterion. A standardized manual scoring criterion for English text discourse coherence is formulated for this paper, and the 1000 English texts in the CELC corpus are scored according to this criterion for discourse coherence.

85-100 points: the overall transition of the English text is natural, the vocabulary is highly articulated and the logic is clear.

70-84 points: the overall transition of the English text is basically natural, the lexical connection is high, and the logic is clear.

50-69 points: the overall transition of the English text is slightly hard, the lexical articulation is low, and the logic is slightly confusing.

0-49: The overall transition of the English text is not natural, the lexical articulation is low, and the logic is confusing.

The English text scores were divided into four grades, and the main scoring criteria were vocabulary articulation and logical coherence. The above 1000 English texts were randomly divided into 500 training sets and 500 test sets, and the 500 English texts in the test sets were scored by the model prediction, and then the scores predicted by the model were compared with the teacher's manual scores, and the scatter of the comparison of the English text discourse coherence scores is shown in Figure 2.

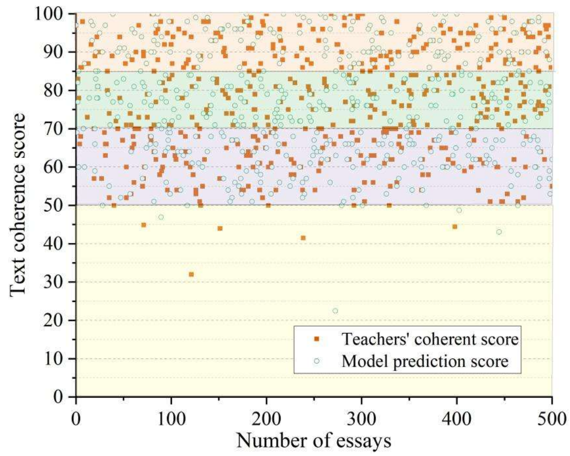

Figure 2: Scatter plot comparison of discourse coherence scores in English texts

In Figure 2, the orange square solid markers indicate the teacher coherence scores, and the green round hollow markers indicate the English text discourse coherence prediction scores of the English grammar analysis model in this paper. Observation of the data distribution shows that the model prediction scores are generally close to the teacher coherence scores, and despite the presence of some data points that differ significantly from the teacher scores, the two show a similar general trend. As discourse coherence is a subjective and abstract evaluation criterion, teachers' ratings may be affected by differences in personal understanding, for example, different personal definitions of coherence, or more harsh or lenient ratings when faced with English texts with many incorrect vocabulary words and difficulties in grasping the author's main idea, so it is reasonable that there are some data points with large differences. From the perspective of quantitative analysis, the average absolute error of this paper's model in scoring the English text for discourse coherence versus the teacher's corrections is 4.103 points, an error value that suggests that the model scores are closer to the teacher's scores. In addition, the Pearson correlation coefficient between the model scores and the teachers' manual scores in this paper reaches 0.789, which is located between 0.6 and 0.8, which indicates that there is a strong correlation between the model scores and the teachers' scores. Taking the above factors into account, it can be concluded that this model has good practical value in practical teaching applications.

IV. Detection of grammatical errors in English texts

Experiments in Chapter 3 show that the grammar analysis model performs well in the coherence diagnosis task. To further explore its application potential in business scenarios, Chapter 4 will focus on the automated detection of grammatical errors, optimize the error correction model with business text characteristics, and solve the grammatical normative challenges in real writing.

IV. A. Experimental preparation

IV. A. 1) Data sets

In order to provide data support for the conduct of the study, this chapter collects and constructs a grammatical error detection dataset for business English texts, and based on the analysis results of the public dataset of the English Grammatical Error Detection Task, applies the characteristics of the types of grammatical errors involved in the public dataset as well as the distribution of the data volume of the different types of errors in the construction of the dataset of grammatical error detection for business English texts. In this section, the public dataset is first analyzed and the conclusion of the analysis is given. Then, according to the analysis results, the collected business English text data are constructed to obey the same distribution of grammatical error detection dataset. Finally, the dataset used for experiments in this chapter is preprocessed.

The experimental data in this chapter is derived from English texts in real business situations, including international business emails, draft business contracts, project reports and marketing texts, etc. A total of 12,000 utterances were collected. All texts were marked by professional business English teachers and covered four types of typical errors: redundant words (e.g., overuse of “very” or redundant modifiers); Inappropriate collocations (e.g., “sign a contract” is misspelled as “sign an agreement”); Missing words (e.g., the qualifier “the” or “shall” is missing from the clause); Word order errors (e.g., inappropriate order of polite phrases in business letters). The total character data for characters marked as correct accounted for more than 90% of the total characters, while the percentage of characters for different grammatical error types was less than 5%, and the amount of data for some of the error categories was even less. This shows that there is a serious data imbalance problem in this task.

IV. A. 2) Experimental setup

This experiment makes the BERT network as a feature representation learning network for the MLM-GEC model. Its network parameters are randomly initialized and the input dimension is set to 1000, the dimension of the hidden layer is set to 1000, the number of network layers is set to 2, and the batch size is set to 1. For the MLM layer, this experiment randomly initializes its state transfer matrix and sets the start and end states to -1E5 to ensure that the state cannot be reached. The model was trained using the SGD optimizer with an initial learning rate of 1E-3 and set the weight decay to 1E-5 with 1000 training cycles.

IV. A. 3) Evaluation indicators and comparison models

This experiment uses PrecisionPrecision, RecallRecall and F1 as evaluation metrics for performance evaluation at three levels.

The performance of this paper's model model was compared with TCN, Bi-LSTM and Bi-LSTM-GEC, all network models using the same experimental setup parameters.

TCN: - In general, time-series data modeling is mostly based on recurrent neural networks and their related variants, such as LSTM, GRU, etc. Since the traditional convolutional neural network CNN does not have the ability to model time-series problems. Until the time-series convolutional network TCN model. This model has been shown to outperform related models of recurrent neural networks on some tasks.

Bi-LSTM: The Bi-LSTM based model obtains the forward and backward hidden states of the input xi at time t. The current hidden state is then obtained by splicing and the output of the model is used as a feature representation of the input. Finally, the probability corresponding to each character in the target state space is obtained by softmax function.

Bi-LSTM-GEC: Based on the structure of the Bi-LSTM model, the Bi-LSTM-GEC model adds a GEC layer to replace the softmax function. A replication mechanism is used to compute the probabilities and predict the output label sequence to obtain the maximum score.

IV. B. Comparative analysis of experimental results

In this section, a multi-dimensional experimental analysis will be conducted based on this framework to reveal the performance advantages of the model in detection, identification and localization tasks, and to verify its ability to solve the data imbalance problem through error class segmentation.

IV. B. 1) Analysis by level

The final experimental results are shown in Table 2. P stands for precision precision and R stands for recall recall.

Table 2: Experimental result/%

Model	Detection level			Recognition level			Location level		
	P	R	F1	P	R	F1	P	R	F1
TCN	46.35	84.38	59.83	40.18	50.78	44.86	7.76	17.58	10.77
Bi-LSTM	43.38	85.95	57.66	41.41	52.05	46.12	8.76	20.84	12.34
Bi-LSTM-GEC	51.28	87.49	64.66	45.86	57.18	50.90	25.83	26.74	26.28
OURS	57.19	89.68	69.84	50.34	60.78	55.07	28.44	33.96	30.96

In the three levels of syntactic error detection (detection, recognition, and localization), the sequence-to-sequence error correction model proposed in this paper, which combines the replication mechanism GEC with the BERT pre-trained language model, leads the way across the board. Its F1 value at the detection level is 69.84%, which is significantly higher than the 64.66% of Bi-LSTM-GEC, the F1 value at the recognition level is 55.07%, which is 4.17 percentage points higher than that of Bi-LSTM-GEC, and the F1 value at the localization level is 30.96%, which is even more than the latter, 4.68 percentage points higher. This indicates that the introduction of the BERT pre-trained language model with the replication mechanism significantly enhances the model's ability to localize complex error types (e.g., word order errors).

IV. B. 2) Experimental analysis by category

In order to further demonstrate the performance of the model, this chapter shows the performance of each type of error (redundant words, improper collocation, missing words and word order errors) at the recognition level, respectively, and the experimental results on each type are shown in Table 3.

Table 3: Experimental results in various categories

Model	Redundant words			Improper matching			Word deficiency			Word order error		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
TCN	30.29	58.46	39.90	16.03	48.81	24.13	22.87	54.99	32.30	6.16	36.7	10.55
Bi-LSTM	32.87	60.93	42.70	18.29	50.32	26.83	25.95	57.67	35.79	7.64	37.29	12.68
Bi-LSTM-GEC	36.96	64.88	47.09	21.67	52.7	30.71	27.62	60.16	37.86	9.22	39.18	14.93
OURS	39.09	67.14	49.41	23.83	54.57	33.17	30.47	63.5	41.18	10.24	42.17	16.48

For the four types of grammatical errors, the F1 values of this paper's model are optimal at the recognition level. For example, the F1 value of redundant word detection is 49.41%, which is 2.32 percentage points higher than that of Bi-LSTM-GEC; the F1 value of improper collocation is 33.17%, which is 2.46 percentage points higher; and the F1 value of word order error is 16.48%, which is 1.55 percentage points higher. This result verifies the effectiveness of the multi-task learning strategy, especially the combination of fine-grained error classification and global semantic characterization, which significantly alleviates the data imbalance problem.

V. Research on the application of automatic assistive platform for business English writing in English writing in colleges and universities

While the previous paper verified the technical effectiveness of the model through grammar error detection experiments, this chapter further embeds this technology into actual teaching scenarios to explore its value in educational practice. The technical validation provides an empirical basis for the platform landing, while the teaching experiment verifies the feasibility of the technical empowerment from the application level, and the two together constitute a complete closed loop from algorithm optimization to educational landing.

The automatic assistance platform for English business writing based on grammar analysis algorithm designed in this paper is applied to actual business English writing in colleges and universities. The two classes of students in a university's class of 2024, majoring in English, who are involved in the course of business English writing are taken as the research objects. The two classes are respectively the experimental class and the control class, and the experimental class and the control class both have 53 students. The experimental class applies the English business writing automatic assistive platform, and the control class adopts the traditional business English writing teaching mode. The experimental class was taught for one semester. The experimental and control classes were directly compared in terms of writing pre-test and post-test scores.

V. A. Pre-laboratory measurements

Prior to the launch of the experiment, a pre-test writing task with voluntary participation but with an emphasis on the need for seriousness was organized for a total of 106 students in two mutually independent parallel classes. The pre-test was designed to assess the students' initial level of business English writing. Subsequently, all students' pre-test scores were accurately entered into a computerized system and processed in depth with the help of a professional data analysis software, SPSS. Through the analysis function of the software, the mean scores and standard deviations of the writing scores of the students in the two classes were calculated, which were used as quantitative indicators to clarify and define the possible differences in writing ability between the two classes. The comparison between the experimental and control groups regarding the pre-test results is shown in Table 4.

Table 4: Comparison of the Pre-test results between two class

Analysis item	Class	M	SD	t	p
Total score of the composition/100	Experimental class	58.87	21.67	0.012	0.923
	Control class	58.69	24.56		
Vocabulary cohesion/25	Experimental class	12.54	7.56	0.081	0.812
	Control class	12.38	7.62		
Logical coherence/25	Experimental class	16.85	6.47	0.122	0.742
	Control class	16.41	5.81		
Grammatical accuracy/25	Experimental class	17.65	8.31	0.-107	0.765
	Control class	17.86	6.99		
Professional Terminology and Business Expression Standardization/25	Experimental class	11.83	4.39	-0.088	0.794
	Control class	12.04	5.67		

There is no significant difference between the experimental group and the control group in the pre-test of business English writing in all the indicators, p-value > 0.05. The total mean score of the experimental group is 58.87 and that of the control group is 58.69, and the difference between the two total scores is only 0.18 points, t=0.012, p=0.923. In the subdimension, the scores of vocabulary articulation of the experimental group of the control group are 12.54 and 12.38, respectively; the scores of logical coherence scores were 16.85 and 16.41; grammatical accuracy scores were 17.65 and 17.86 and terminology standardization scores were 11.83 and 12.04, respectively, and the difference in the mean values of all four did not reach statistical significance, p>0.74, indicating that there was homogeneity in the initial writing ability of the two groups of students.

V. B. Post-experimental tests

While the previous section established the homogeneity of the experimental and control groups through the pre-test data, this section relies on the comparison of the post-test results to quantitatively present the significant effect of the assistive platform on the improvement of writing ability, and the rigor of the pre and post-test design provides a double guarantee of the reliability of the conclusions.

V. B. 1) Comparison of experimental and control groups regarding posttest results

After a semester of business English writing teaching practice under different teaching modes, the experimental and control group students were then tested on relevant business English writing, and the comparison of the experimental and control groups about the results of the post-test is shown in Table 5.

Table 5: Comparison of the test results between two class after the experience

Analysis item	Class	M	SD	t	p
Total score of the composition/100	Experimental class	90.41	7.48	14.283	0.000**
	Control class	71.01	18.83		
Vocabulary cohesion/25	Experimental class	23.16	1.48	8.793	0.000**
	Control class	18.14	4.82		
Logical coherence/25	Experimental class	22.49	2.06	7.012	0.000**
	Control class	19.06	5.08		
Grammatical accuracy/25	Experimental class	24.01	0.74	11.572	0.000**
	Control class	17.03	5.63		
Professional Terminology and Business Expression Standardization/25	Experimental class	20.75	4.15	9.272	0.000**
	Control class	16.78	7.87		

Table 6: Comparison of the pre - and post-test analysis results in the experimental class

Analysis item	Class	M	SD	t	p
Total score of the composition/100	Pre-test	58.87	21.67	-23.348	0.000**
	Post-test	90.41	7.48		
Vocabulary cohesion/25	Pre-test	12.54	7.56	-17.226	0.000**
	Post-test	23.16	1.48		
Logical coherence/25	Pre-test	16.85	6.47	-13.201	0.000**
	Post-test	22.49	2.06		
Grammatical accuracy/25	Pre-test	17.65	8.31	-14.612	0.000**
	Post-test	24.01	0.74		
Professional Terminology and Business Expression Standardization/25	Pre-test	11.83	4.39	-18.079	0.000**
	Post-test	20.75	4.15		

After one semester of teaching experiment, the total mean score of the experimental group using the assistive platform was significantly higher than that of the control group at 90.41, and the total score of business English writing of the control class of the experimental traditional teaching mode was 71.01, t=14.283, p=0.000. In the subdivided dimensions, the vocabulary articulation scores of the two classes were 23.16 and 18.14, respectively, t=8.793; the logical coherence scores were 22.49 and 19.06, t=7.012; grammatical accuracy of 24.01 and 17.03, t=11.572; and terminology standardization of 20.75 and 16.78, t=9.272, the scores of the four dimensions were significantly better than those of the control group, p=0.000. Meanwhile, the standard deviation of the experimental group was significantly narrowed down, which indicated that the level of their writing was more consistent.

V. B. 2) Comparison of pre and post-test analysis results in the experimental group

The results of the pre and post-tests of the experimental group were analyzed in a longitudinal comparison, and the comparison of the results of the pre and post-test analyses of the experimental class is shown in Table 6.

Comparison of the experimental group's own pre and post-tests showed that the total score improved from 58.87 to 90.41, t=-23.348, p=0.000, an improvement of 53.6%. The subdimensions improved significantly: lexical articulation from 12.54 to 23.16, t=-17.226; logical coherence from 16.85 to 22.49, t=-13.201; grammatical accuracy from 17.65 to 24.01, t=-14.612; and terminology regularity from 11.83 to 20.75, t=- 18.079, with p-values less than 0.001. And the standard deviation of the posttest of the experimental group was substantially reduced, indicating that the auxiliary platform effectively narrowed the differences in the business English writing ability of college students.

VI. Conclusion

In this study, through the deep integration of grammar analysis algorithm and deep learning, we constructed an automated assistance platform for English business writing, and achieved the following core results:

The CRF-RST-Style model has an MAE of 2.071 and a PCC of 0.702 in lexical articulation diagnosis, which is significantly better than the traditional model, verifying the effectiveness of context-dependent parsing.The error correction model fusing GloVe word vectors and BERT has an F1 value of 69.84% in the task of grammatical error detection, and the localization error F1 value is improved by 4.68 percentage points, solving the the challenge of recognizing complex errors in business texts.

The PRA of the model on the Accident and Earthquake datasets reaches 96.57% and 97.98% respectively, indicating that the global semantic-driven strategy can effectively capture long-distance logical dependencies and outperforms traditional methods such as HMM with 90.24%.

In the teaching experiment, the total posttest score of the experimental group applying the English writing automation assistance platform is 90.41, which is 27.3% higher than the 71.01 of the control group in the traditional teaching mode, with grammatical accuracy close to the full score (24.01/25) and the standard deviation reduced by 65.5%, which confirms that the platform significantly improves the writing standardization and realizes personalized assistance.

References

[1] Magiman, M. S. A. (2022). Enhancing students critical thinking skills in writing by promoting ESP-based language learning environment. Journal of Positive School Psychology, 6(4), 3717-3730.

[2] Srinivasan, S., Thangaraj, R., & Mathew, J. (2024). Role of experiential learning program on business writing skills of management students. Business and Professional Communication Quarterly, 23294906241228244.

[3] Atique, S. S., & Khan, I. (2015). The Writing needs of Business Students: A Teacher’s Perception in an EFL context. Journal of Education & Social Sciences, 3(2), 231-244.

[4] Lama, A., & Suhodolli, M. (2024). Challenges in mastering academic writing: a case study of English language learners at the university for business and technology. Edelweiss Applied Science and Technology, 8(6), 84-99.

[5] Ria, T. N., & Malik, D. (2020, November). Syllabus design in Business English based on the needs of Economics students. In ELT Forum: Journal of English Language Teaching (Vol. 9, No. 2, pp. 140-149).

[6] Masyhudianti, U. K., Sutomo, N., & Suparno, S. (2018). The Effectiveness of Schoology to Teach Writing Viewed from Students' Creativity. International Online Journal of Education and Teaching, 5(4), 943-955.

[7] Farooq, M. S., Uzair-Ul-Hassan, M., & Wahid, S. (2020). Opinion of second language learners about writing difficulties in English language. South Asian Studies, 27(1).

[8] Lin, C. J., Hwang, G. J., Fu, Q. K., & Chen, J. F. (2018). A flipped contextual game-based learning approach to enhancing EFL students’ English business writing performance and reflective behaviors. Journal of Educational Technology & Society, 21(3), 117-131.

[9] Al-Mutawa, A. S., Al-Kandari, H. S., & Fayez, F. M. (2024). An Analysis of Arab Undergraduate Students' Writing Performance: Applying SWOT Framework. Journal of Language Teaching and Research, 15(2), 436-447.

[10] Li, X. (2018). Influence of computer-aided instruction model on business English writing teaching effect. International Journal of Emerging Technologies in Learning (Online), 13(3), 197.

[11] Cardon, P., Fleischmann, C., Aritz, J., Logemann, M., & Heidewald, J. (2023). The challenges and opportunities of AI-assisted writing: Developing AI literacy for the AI age. Business and Professional Communication Quarterly, 86(3), 257-295.

[12] Fitria, T. N. (2021). Grammarly as AI-powered English writing assistant: Students’ alternative for writing English. Metathesis: Journal of English Language, Literature, and Teaching, 5(1), 65-78.

[13] Frankenberg-Garcia, A., Lew, R., Roberts, J. C., Rees, G. P., & Sharma, N. (2019). Developing a writing assistant to help EAP writers with collocations in real time. ReCALL, 31(1), 23-39.

[14] Alharbi, W. (2023). AI in the foreign language classroom: A pedagogical overview of automated writing assistance tools. Education Research International, 2023(1), 4253331.

[15] Nazari, N., Shabbir, M. S., & Setiawan, R. (2021). Application of Artificial Intelligence powered digital writing assistant in higher education: randomized controlled trial. Heliyon, 7(5).
[16] Lee, M., Gero, K. I., Chung, J. J. Y., Shum, S. B., Raheja, V., Shen, H., ... & Siangliulue, P. (2024, May). A design space for intelligent and interactive writing assistants. In Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (pp. 1-35).