Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2025 SCPE. Volume 26, Issues 3, pp. 1348–1362, DOI 10.12694/scpe.v26i3.4314

EXPLORING A NEW MODEL OF COLLEGE ENGLISH TRANSLATION CLASSROOM VIA NATURAL LANGUAGE PROCESSING AND COMMUNICATION TECHNOLOGY

YUCHEN GUO*

Abstract. The crucial duty of developing translation skills for China's modernization falls on higher education institutions that teach translation. Information and intelligent technology are becoming increasingly ingrained in people's lives as civilization grows and develops. In this work, we use natural language processing and communication technologies to build a new type of university English translation classroom. To address the challenge of inferring semantic implication linkages in natural language processing, we put forth a deep learning model based on semantic rounding and semantic fusing. The technique can be applied to university translation classes to help basic translation tasks with effective reading comprehension. Furthermore, we developed a wireless classroom interaction system that enables effective interoperability between teachers and students in the classroom by embedding a natural language processing model in real time. Our natural language processing model performs exceptionally well and is capable of making predictions in real time, according to experimental results. The entire solution gives universities English translation classes a whole new experience.

Key words: Higher education,English translation, Interaction system, Wireless communication, Natural language processing model

1. Introduction. Classroom interactive systems have drawn increasing attention as a crucial tool to support classroom information interaction and assist teachers in understanding students' learning status in real time, given the ongoing development of education informatization and the ongoing transformation of traditional teaching methods [1]. As a significant area of study in computer science and artificial intelligence, natural language processing (NLP) has emerged in recent years. Its goal is to develop ideas and techniques that would enable humans and computers to communicate naturally. In addition to being a branch of computer science, natural language processing incorporates knowledge from other academic fields including linguistics and mathematics [2]. Although it focuses on human language in everyday conversation, its study is essentially distinct from that of traditional human linguistics. Instead of studying human language per se, natural language processing focuses on creating computer systems—particularly software systems—that enable efficient natural language communication between people and machines [3].

When it comes to teaching translation and classroom interaction systems, natural language processing is especially crucial. Teaching translation requires the capacity to comprehend and process material, and NLP offers several useful tools to help with this process. A "fill-in-the-blank" question, for instance, is comparable to a Cloze-style query in that the computer reads and comprehends the text, then extracts words or entities from the sentences and provides an answer in accordance with the query [4, 5]. Conventional models often encode the question and document, output the answer, or iteratively update the multilayer network's attention mechanism's focus of attention, finally producing the answer consisting of the words that have received the greatest attention. These methods, nevertheless, frequently overlook the larger context in favor of concentrating solely on a few chosen phrases [6]. This chapter focuses on the functions of semantic rounding networks and semantically aligned fusion methods in the construction of a new deep learning model, SDF-NN. We base this on the proposal of semantic rounding networks and semantic fusion methods. These two significant enhancements increase the accuracy of the model, encourage the complete fusion of local inference results, and lessen the detrimental effects of interfering semantics on the final prediction outcomes [7, 8].

Furthermore, we created a wireless classroom interaction system that supports real-time teacher-student interaction through the natural language processing model integration. In this way, natural language processing

*School of Foreign Languages Fuyang Normal University, Fuyang Anhui 236000, China (jiliangxie@jou.edu.cn)

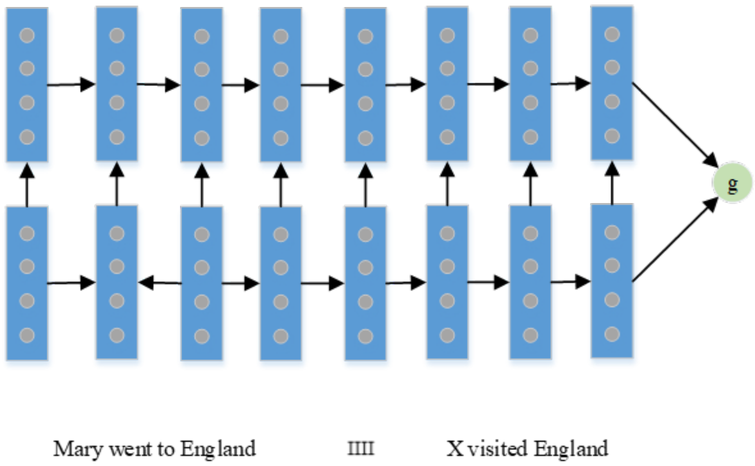

Fig. 2.1: Deep LSTM Reader Model.

strengthens the case for the modernized educational model by improving classroom interactions while also increasing the effectiveness of translation instruction.

2. Related work.

2.1. Natural Language Processing Based On Deep Learning.

1. Natural language understanding and in-depth education. An overview of the evolution of Natural Language Processing (NLP) with deep learning may be found in this section [9, 10]. It goes into great length on how machine learning challenges can be derived from problems with natural language understanding, including the requirement to convert computer-readable mathematical symbols from human-readable characters.

2. Inserting Words. The notion of word embedding and its significance in deep learning can be thoroughly explained in this section. First, the original word representation—one-hot representation—as well as its drawbacks—particularly the lexical gap phenomenon—are covered. After that, the idea of distributed representation (DR) is presented in order to clarify how it improves upon the One-hot representation's drawbacks [11, 12].

3. Vector bases of pre-trained words. Word vector libraries that have already been trained, like GloVe, can be covered in this area along with their uses and benefits for various NLP tasks (including named entity identification, word analogies, and word similarity) [13, 14].

Readers will be better able to comprehend the essence of each topic and its place in your research, and the sections on related work will be more logically organized.

2.2. Attentional Models. The first deep learning model that has been proposed, called Deep LSTM Reader, is a very basic model that encodes a query and a document independently using a two-layer LSTM, and then classifies data based on the representation that has been created from the two layers. This extremely basic model, which solely relies on encoding, misses the relevant characteristics of documents and queries. (Fig. 2.1).

The report then presented two models that Google DeepMind has developed: the eager reader model and the attentive reader model. First, the document and query are represented separately in the attention reader model. The query is encoded using bidirectional long short term memory (LSTM) and its overall representation is obtained by splicing the outputs of the forward and reverse last hidden layers. Similarly, the document is also encoded using bidirectional LSTM and its encoded representation of each lexical element (i.e., word in the document) is obtained by splicing the corresponding forward and reverse hidden layers. Finally, the overall representation of the document is a weighted average of all the lexical elements in the document, where the

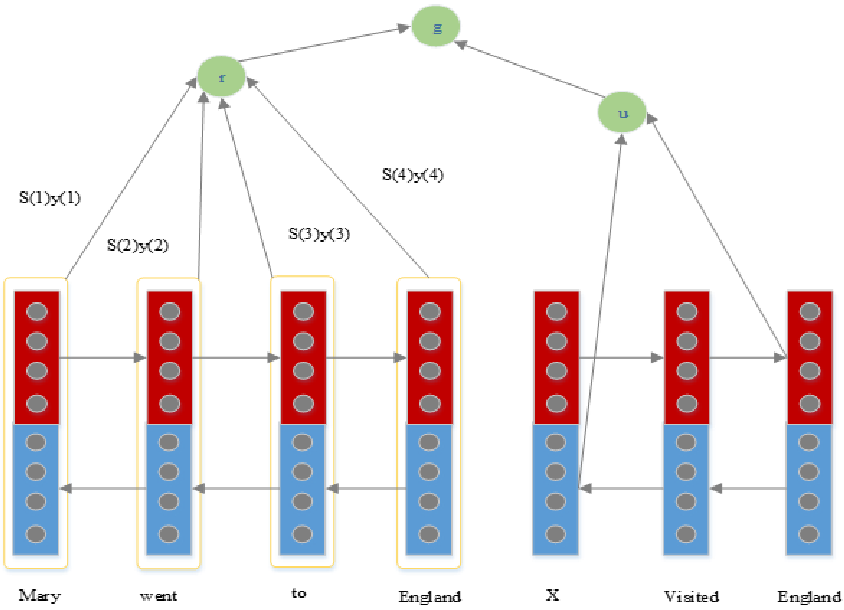

Fig. 2.2: Attentive reader model.

weights are determined by the generated attention, and the weights indicate the importance of the corresponding lexical elements in answering the query.

These models have a direct connection to the classroom interaction system and model you have suggested. Similar attentional mechanisms might be employed in the system you are building to gauge how important various text segments are in responding to students' inquiries and maximizing teacher-student engagement. Additionally, when handling bidirectional information flow in classroom interactions, the deployment of bidirectional LSTMs may also be instructive. This implies that important technological components from these linked works may immediately offer your suggested system technical help and inspiration.

Then the documents and query representations are used for classification [15, 16]. This is shown in Fig. 2.2. The Impatient Reader has been improved in that instead of encoding queries as a whole as in the Attentive Reader model, the token of each query is related to the token of the document. This mechanism is similar to reading each token in a query and then focusing on the information of the corresponding token in the document [17]. This model has a more complex attention mechanism, but it may not be effective, because in terms of the actual human reading comprehension mindset, it is impossible to read a word in the query and then go back to read the original text again when answering the question, which is too inefficient, and long documents may also affect memory. The structure of the model is shown in Fig. 2.3.

The Stanford AR model mainly uses a bidirectional LSTM to encode the document and query separately, and uses the correlation between the words and the query to obtain the Attention value, which is used to weight the embedding of the document to obtain a final output vector for answer prediction. The model structure is shown in Fig. 2.4.

The Attention Sum Reader model obtains the associated representation vectors by encoding the document and the question, respectively, using two bidirectional GRUs. The outcome, which can be seen as the attention matrix, can be thought of as the weight of each token in the document in relation to the query. Ultimately, each token's likelihood in the text is normalized using the softmax function, and the result with the highest probability is regarded as the answer to the question. That's what the figure depicts. The AOA model is suggested because Q&A should be predicated on the mutual attention of the inquiry and the document, whereas the aforementioned models are based on one-way attention.

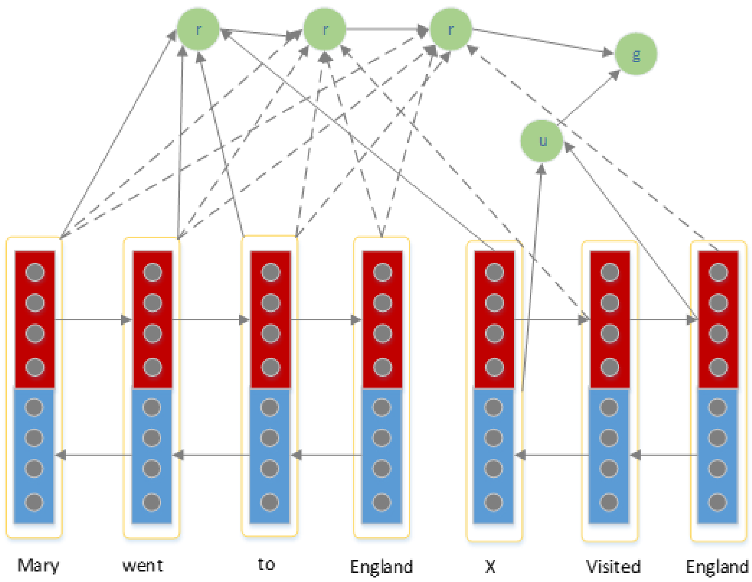

Fig. 2.3: The model of the patient reader.

2.3. Classroom Interaction. The so-called intelligent classroom is actually a multimedia classroom that operates and controls audio-visual equipment, computers, projectors, light, electricity and other devices in the classroom, facilitates access to teaching resources and teaching activities for teachers and students, and provides information storage and real-time feedback [18]. Another development idea of classroom interactive system is to modify and upgrade the existing classroom hardware in a limited way. Both of these ideas can facilitate interaction between teachers and students, but they also have many differences [19]. There are many educational research institutions and companies in China and abroad that are focusing on smart classrooms, including McGill University, the University of Chicago, DELL, and Intel Corporation. For example, DELL has proposed an intelligent classroom solution with the goal of creating an interactive and collaborative learning environment.

In China, some companies have also launched intelligent classroom solutions, such as the Xunjie II intelligent classroom developed by Shanghai Excellence Electronics Co. The intelligent classroom mainly consists of the following components: control panel, fully automatic guide, multimedia external devices, teacher and student scene cameras, etc. Users can choose student interaction modules according to actual needs, as shown in Fig. 2.5.

At present, research on smart classrooms has made some progress in both theoretical and applied research, and these devices have met the needs of teachers and students for daily interaction to a certain extent [20].But these gadgets necessitate a significant hardware upgrade for the classroom, which is expensive in terms of engineering, time, and renovation expenses. Upgrading any component of the system later on is likewise challenging. An alternative perspective on the teacher-student interaction system is comparable to classroom voting devices like the SunVote Conference Voting Device S52Plus, which can communicate at a distance of approximately 30 meters, uses 2.4G frequency wireless radio frequency technology, is moderately sized, and uses CR2032 coin cell batteries. These devices are used by many college students in North America. It is mostly utilized for staff training, yearly meetings, product roadshows, academic conferences, quality assessment and evaluation, etc.

3. Methodology.

3.1. Semantic Entailment Relation Inference Model Based On Semantic Discarding And Fusion. The relationship between the two sentences is incorrectly anticipated to be contradictory if the strongest

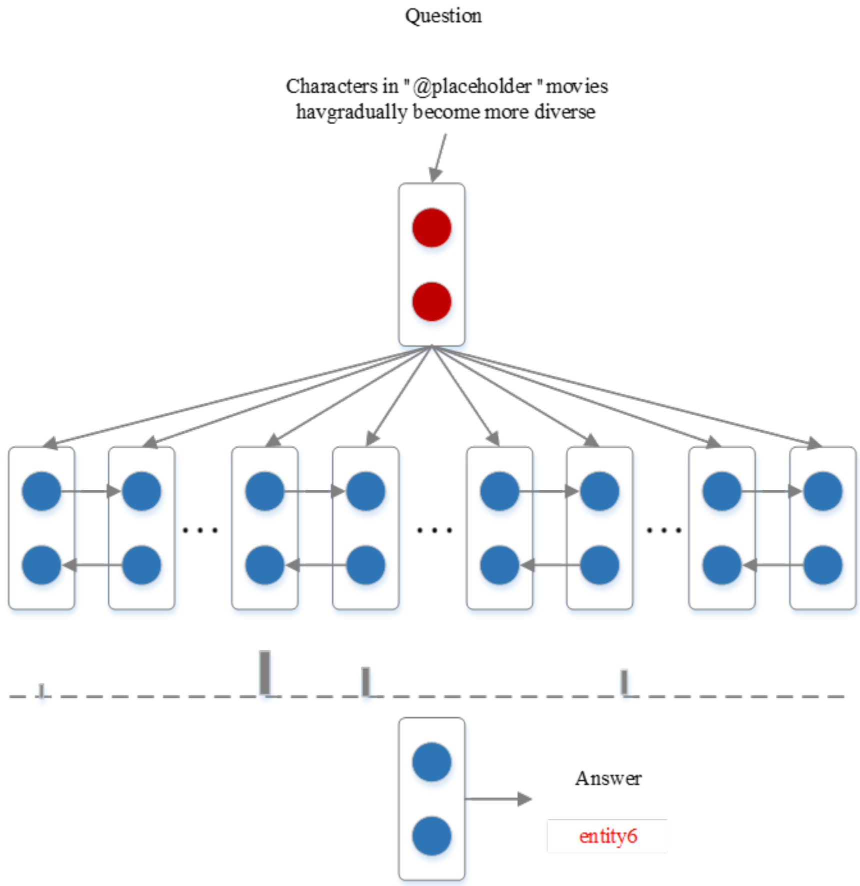

Fig. 2.4: Stanford ar model.

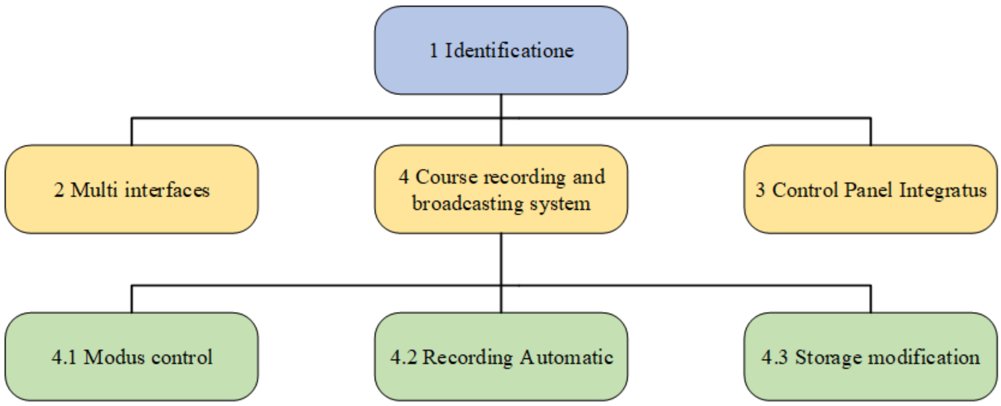

Fig. 2.5: Functional block diagram of excellent intelligent classroom.

semantic relation is utilized as the final prediction. To predict the proper result as "neutral," the model needs to take into account the combined findings of many local inferences. Therefore, in order to rationally fuse all

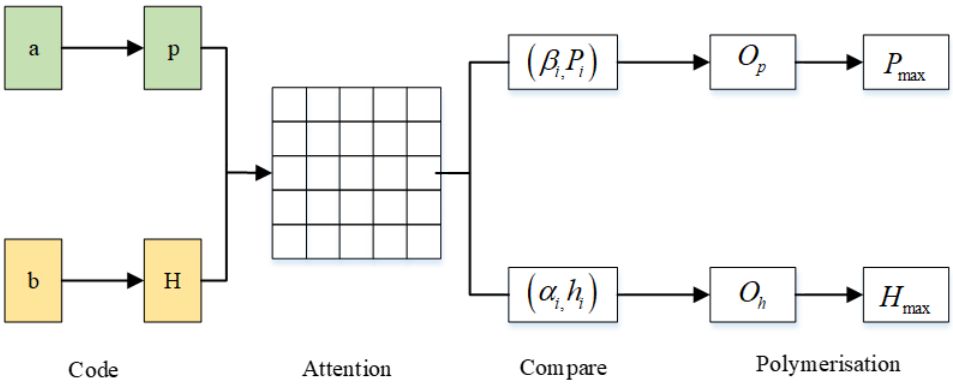

Fig. 3.1: Network structure of model based on decomposed attention.

local inference outcomes, this chapter suggests a semantic fusion alignment approach. Drawing from the aforementioned constraints of the two earlier models, the first step involves designing a semantic discarding network (SDN) to exclude extraneous or disruptive semantic information during the "comparison" phase. While all extracted semantic information is passed forward in this network, some semantic information is purposefully dropped during training.

In previous deep learning models based on decomposing attention, four processing steps are generally summarized: coding, attention mechanism, comparison, and aggregation. This is shown in Fig. 3.1.

Assume two sentences a and b, a is an Embedding representation of the premise of length m and b is a word vector representation of the hypothesis of length n. After encoding the two sentences using the encoder, the encoded representation of the word vector is obtained as the matrix $P = [P_1, \ldots, Pm], \forall i \in [1, \ldots m]$ and $H = [h_1, \ldots, hn], \forall i \in [1, \ldots n]$. The corresponding aligned text pairs are then obtained by applying the corresponding decomposition attention mechanism to the attention matrix. (α_j, h_j) and $(\beta_i, p_i)$are called aligned text pairs, where α_j is a subfragment in P that is aligned withh_j , which is a recoded representation of P based on the attention distribution of h_j to each word in P . β_i is a subfragment in H that is aligned with p_i, which is a recoded representation of H based on the attention distribution of pi to each word in H . This is also known as aligning related segments based on attention and highlighting salient features. The above steps are designed to extract semantic features from the sentences. In the "comparison" phase, a single feedforward neural network or LS TM network is usually used to extract the semantic features between aligned text pairs and obtain the corresponding local inference results . The aligned text pairs are fed into the feedforward neural network G, and the local inference result O is obtained as shown in Fig. 3.2. However, in natural language, different aligned text pairs have different relationships, and the internal feature relationships are not consistent, which means different local inference results. For example, "wears, dressed in" means the same thing, and "in the morning, at night" means the opposite. Therefore, different functions should be applied to extract the different relationships between the aligned text pairs, as shown in Fig. 3.3. Each aligned text pair is passed through k different feedforward neural networks G. Through the learning of these feedforward neural networks, multiple local features are generated for each aligned text pair. The gate function g is then defined to determine the weight of each G function, i.e., each local inference result, and finally weighted and processed so that each pair of aligned text pairs still produces a corresponding local inference result. This inference result will be more accurate.

As analyzed in the previous section, these traditional networks extract and analyze all the extracted semantic information, which includes interfering semantics that can have a negative impact on the final model. Therefore, to address this problem, we propose a feedforward discard network (SDN) that differs from the traditional approach, as shown in Fig. 3.4.

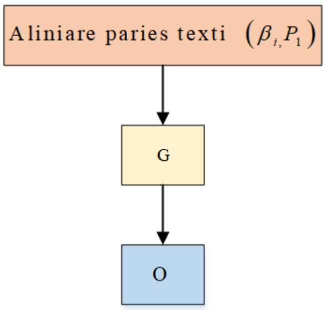

Fig. 3.2: Local inference results obtained through a feedforward neural network.

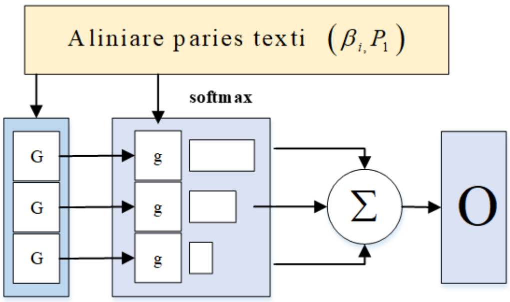

Fig. 3.3: Local inference results obtained through multiple feedforward neural networks.

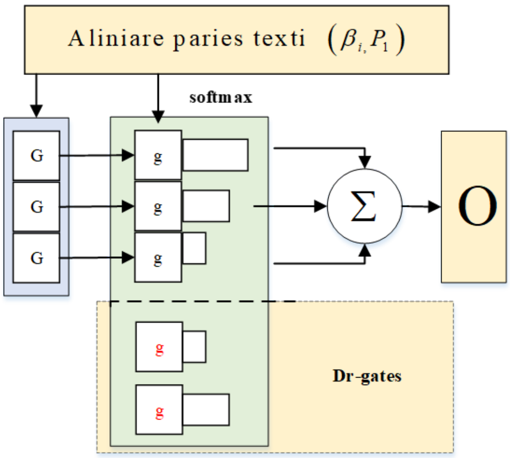

Fig. 3.4: SDN Structure.

3.2. Semantic Fusion Alignment Methods. In the traditional approach, the obtained local inference results are aggregated directly by maximum pooling or average pooling. This is shown in Fig. 3.5.

The obtained local inference results are directly processed by maximum pooling to the most significant features, and then fed into the softmax layer for the final prediction. This simple treatment, as analyzed above, ignores the relationship between the local inference results and does not include their combined decision information. Each column of the local inference outcome matrices Op and Oh is a local inference outcome, and

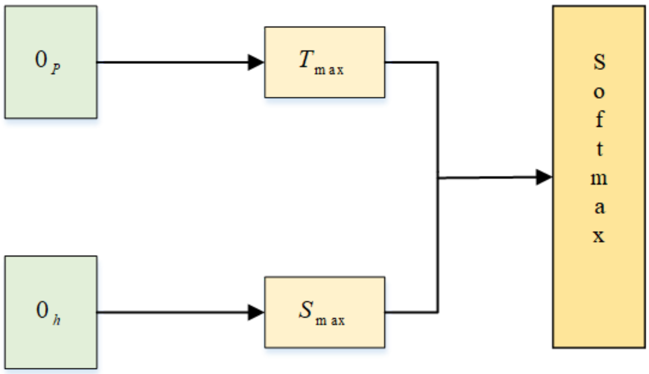

Fig. 3.5: Processing local inference results with max pooling.

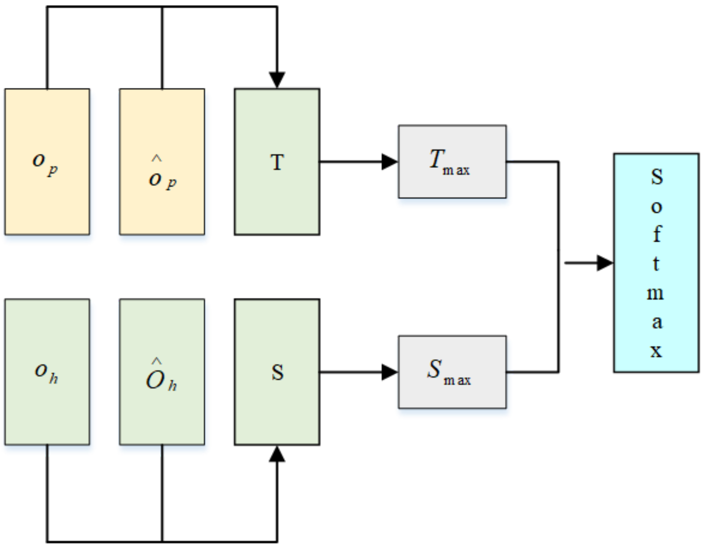

Fig. 3.6: Structure of SFA.

each row of their corresponding transpose matrices represents a local inference outcome.First we multiply the local inference result matrix with its transpose matrix (see Fig. 3.6):

$$U = O_p^T O_p \tag{3.1}$$

$$V = O_h^T O_h \tag{3.2}$$

U11 in the U matrix represents the attention of the first local inference result in O_p to itself, U21 represents the attention of the first local inference result in O_p to the second inference result, and the same for the remaining elements. each element of the V matrix is also aligned with the elements of the U matrix. The main diagonal elements of the resulting alignment matrices U and V are set to 0, because the elements should not be aligned with themselves and the attention of the local result on itself needs to be eliminated.

$$u_i = (U_i) \tag{3.3}$$

$$v_j = (V_j) \tag{3.4}$$

$$\hat{O}_p^i = O_p \cdot u_i, \forall i \in [1, \ldots, m] \tag{3.5}$$

$$\hat{O}_h^i = O_h \cdot v_j, \forall j \in [1, \ldots, n] \tag{3.6}$$

The elements in the U matrix are normalized by row. After normalization, the first column u1 of the matrix represents the attention weight distribution of the first partial inference result on the other inference results in O_p. The second column u2 represents the attention weight distribution of the second partial inference result on the other inference results in O_p . The second column u2 represents the attention weight distribution of the second partial inference result on the other inference results in O_p , and so on. After normalization, the first column of the matrix, v1, represents the attention weight distribution of the first inference result to the other inference results, and the rest is the same.

3.3. Overview of GIS technology and its related applications. We present a comprehensive inference model for semantic entailment relations based on the decomposed attention mechanism, SDF-NN (Semantic Dropping and Fusion Neural Network), based on the previous findings and the two novel networks and methodologies. We present a comprehensive inference model for semantic implication relations called SDF-NN (Semantic Dropping and Fusion Neural Network), which is based on the decompositional attention mechanism. This model can fully illustrate the efficacy of the semantic fusion alignment (SFA) and semantic dropping network (SDN) techniques. The model has the same four steps as the previous decomposition-based attention model: coding, attention, comparison, and aggregation. The overall framework of the model is shown in Fig. 3.7. Assume that two sentences$a = (a_{1,....,}a_m)$and $b = (b_{1,....,}b_n)$ a is the word vector representation of the premise of length m (Word Embedding representation) and b is the word vector representation of the hypothesis of length n. We assume that a_i , $b_i \in Rd$ and d is the dimension of the word vector. The goal is to predict the label y to determine the relationship between a and b. y can be Neutral, Contradiction, or Entailment. (1) Sentence encoding First, two sentences are input into the bidirectional LS T M in temporal order, and the output of each hidden layer is used as the encoded representation of the word vector of the corresponding input.

$$p_i = biLSTM\,(a_i) \tag{3.7}$$

$$h_j = biLSTM\,(b_j) \tag{3.8}$$

Matrix $p = [p1, \ldots, pm] \in R2r \times m$ and $H = [h1, \ldots, hn] \in R2r \times n$ is the output of boils T M hidden layer. r is the number of neurons in the hidden layer. the LSTM model is formulated as follows.

$$i_t = \sigma\,(W_i x_t + U_i h_{t-1}) \tag{3.9}$$

$$f_t = \sigma\,(W_f x_t + U_f h_{t-1}) \tag{3.10}$$

$$o_t = \sigma\,(W_o x_t + U_o h_{t-1}) \tag{3.11}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tanh\,(W_c x_t + U_c h_{t-1}) \tag{3.12}$$

$$h_t = o_t \tanh\,(c_t) \tag{3.13}$$

$a_i$and b_j are entered into the model sequentially according to the time series t, respectively, as x_t , LSTM utilizes a memory version, including an input gate i_t , a forget gate ft , an output gateo_t , and a memory unit c_t , to generate a hidden layer output h_t . Also, using $biLSTM$ encoding is more efficient than $LSTM$ encoding. It is to encode the sentence forward and then encode it backward, and each word will have two corresponding hidden layer outputs, which will be concatenated as the final encoded form of the word.

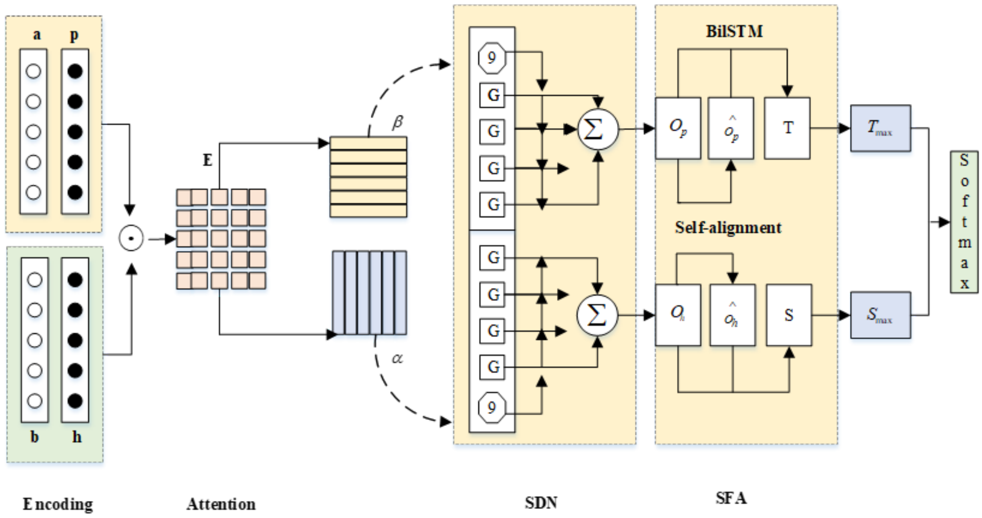

Fig. 3.7: SDF-NN Structure Diagram.

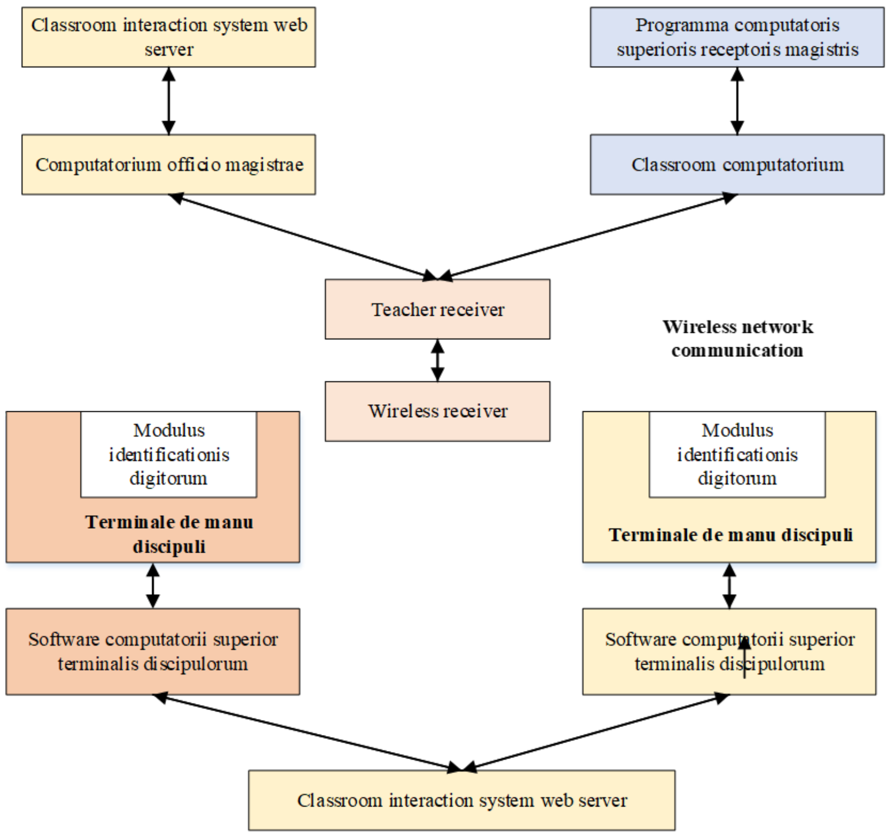

Fig. 3.8: Schematic diagram of classroom interaction system based on wireless communication.

3.4. Functional Modules Of The Wireless Communication-Based Classroom Interactive System. Based on the needs of most users, this wireless communication-based classroom interaction system is designed. The specific structure of the wireless communication-based classroom interactive system is shown in Fig. 3.8.

The teacher receiver host software is installed on the classroom computer, and the teacher receiver and wireless receiver are fixed and physically connected in the classroom. To attend class, each teacher merely needs to place the unique SD card into the teacher terminal. The instructor just needs to bring their SD card with them after class to use the office card reader to transmit the data from the SD card to the web server. The main functions of the classroom interactive system are: (1) The webmaster first adds class and student and teacher user information. (2) The student user logs into the web page and connects the student handheld terminal to the web page through the student handheld terminal host software. Then the MAC address serial number of the handheld terminal chip is passed into the front-end of the webpage, followed by the fingerprint registration of the student handheld terminal, and the address serial number is bound to the user name after the registration is successful. If the user has already registered the fingerprint or the student's handheld terminal has been registered by other students, the above operation cannot be performed.(3) The teacher downloads the list of students and their corresponding MAC address numbers of the class he/she wants to teach from the web page and puts them into the SD card for the teacher's receiver. The SD card is then inserted into the Teacher Receiver during class. The software on the teacher's receiver has various function buttons to enable the class roll call, class quiz ABCD questions or Y/N questions, collect the students' answers and display them on the projector, and then store them on the SD card on the teacher's receiver. (4) Finally, the information stored in the SD card can be stored in the web server for subsequent viewing and processing, so that we can have a more comprehensive understanding of the actual learning status of each student.

4. Experiments.

4.1. Dataset And Parameter Settings. The model was assessed using the Stanford Natural Language Inference (SNLI) dataset, which is made available to the public. The 570k English sentence pairings in the dataset have been manually labeled by several people. Three sets make up the dataset: a test set, a validation set, and a training set. The model is trained using the training set, verified using the validation set, monitored during training to avoid overfitting, then tested once more using the test set to assess the model's performance. As in previous work, we remove the samples labeled with ”-” (the ”rest” class) from the dataset, leaving 549,367 sentence pairs in the training set, 9842 sentence pairs in the validation set, and 3,632 sentence pairs in the test set. The remaining samples in the training set are 549,367 sentence pairs, the remaining samples in the validation set are 9842 sentence pairs, and the remaining samples in the test set are 9824 sentence pairs. In this model, the dropout of the feedforward neural network in the SDN is set to 0.2, and the dropout of the rest of the model is set to 0.3. Set the training batch size to 128, i.e., the number of samples for one input model training is 128, and the loss function for this training is the average of these 128 samples. Depending on the performance of the training machine, this value can be adjusted accordingly. In the SDN layer, the final experimental model is set up with 3 G-functions and 5 gate functions, which can be tuned in more detail to achieve better performance. The training loss function is a multi-class cross-entropy function, and the optimization method is Adam (Adaptive Moment Estimation), which has the advantage that after bias correction, the learning rate of each iteration has a fixed range, making the parameter correction is relatively smooth. The whole model is implemented based on TensorFlow, a second-generation artificial intelligence learning system developed by Google based on DistBelief.

4.2. Performance Analysis. Table 4.1 shows the accuracy comparison of the SDF-NN model and some related models trained and tested on the SNLI dataset. Para is the number of parameters, the first row of the table is a classifier-based feature extraction model, which is considered as a benchmark comparison for the semantic implication inference problem. The next set of models (2) to (3) are based on sentence encoding. The third group of models (4) to (7) are based on attentional mechanisms. The last group of (9) and (11) models are integration models. It can be seen that the proposed model, SDF-NN, achieves the highest accuracy of 88.2% in a single model. At the end of the table, an elimination analysis is also performed to show the impact of two key design modules of the model on the overall model performance. We first remove the drop gates (dr-gates)

Table 4.1: Performance Comparison of Models on SNL Datasets.

Models	Para	Train(%)	Test(%)
(I)Unigram and bigram features [Bowman et al. 2015]	-	99.7	78.2
(2)300D LSTM encoders [Bowman et al, 2016]	3.0M	83.9	80.6
(3)300D Tree-based CNN encoders [Mou et al, 2015]	3.5M	83.3	82.1
(4)100D word-by-word attention [Rocktaschel et al, 2015]	250K	85.3	83.5
(5)600D BILSTM with intra-attemton [Liu et al. 2016]	2.8K	85.9	85.0
(6)200D decomposable attention models[Pankh et al, 2016]	580K	90.5	86.8
(7)300D re-read LSTM [Sha et al- 2016]	2.0K	90.7	87.5
(8)60OD ESIM[Chen et al. 2016]	4.3M	92.6	88.0
(9)600D ESIM+ Syntactic(Ensemble)	7.7M	93.5	88.6
(10)BIMPM [Wang et al, 2017]	-	-	88.9
(11)BIMPM(Ensemble)	6.4M	93.2	88.8
5OOD SDF	6.4M	92.8	88.2
5OOD SDF w/o dr-gates (SDN)	6.3M	91.0	87.7
500D SDF w/o SDNS	5.3M	90.3	87.5
5OOD SDF w/o SDN+SFA	1.5M	90.1	87.0

Table 4.2: Decomposition accuracy of SDF-NN model.

Models	N(%)	E(%)	C(%)
(Bowman et al. 2016)	80.6	88.2	85.5
(Wang ang Jiang 2015)	81.6	91.6	87.4
(Parikh et al. 2016)	83.7	92.1	86.7
SDF-NN(ours)	84.3	92.0	88.1

from the SDN network and the accuracy drops to 87.7%. Then we remove the entire SDN network and replace it with a simple feedforward neural network, and the accuracy drops to 87.5%. Finally, the SFA is removed and a simple maximum pooling method is used to handle those local inference results, and the accuracy drops to 87.0%. As can be seen, several key parts of the model are designed to have a very positive effect on the performance of the model.

Table 4.2 presents the accuracy results of the model for each of the three categories tested in the validation set of the SNLI dataset. It can be seen that the overall accuracy of the model is mainly due to the ”implicit” category, while the main accuracy loss is in the ”neutral” category. The reason for this may be that for the ”implicit” category, it is beneficial to discard the distracting information and consider the relationship between the segments globally for the final inference. However, for the ”neutral” category, there may not necessarily be a direct correlation between the segments of the statement, and forcing some segments to be aligned when decomposing attention may have a negative impact on the final result.

There are two very important parameters in the SDF-NN model, the number of functions G x and the number of discarded gates (drgates) y. Since the weight of all gate functions sums to 1, it is considered that the larger the number of y, the higher the weight of discards, i.e., the more information is discarded in the model. Fig. 4.1 shows the accuracy using different settings of the x and y parameters. Two patterns emerge: (a) First, we fix y to 2 and increase x from 0, and the accuracy rate starts to increase and then level off. This is because the function G serves to fully extract features from multiple aligned text pairs, i.e., local inference results, and we need enough functions G to extract feature information from the data, but this extraction ability will level off as the function G continues to increase. In this case, x is best set to 3. (b) Then, we fix x at 3 and increase y from 0 to 4. The model achieves the best performance for y = 2. This reflects that discarding information has a positive effect on the performance of the model, but too much information is discarded as y increases, which reduces the performance of the model. From these two sets of experiments, and based on the consideration

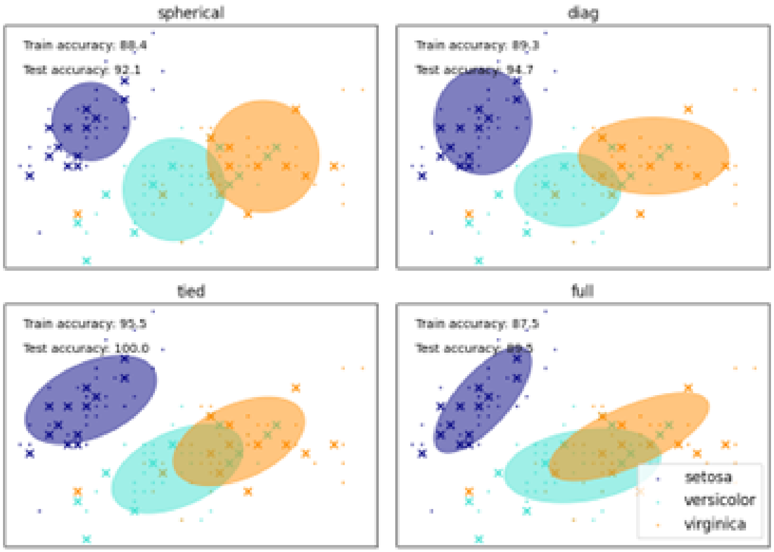

Fig. 4.1: Analysis of key parameters of SDF-NN.

Table 4.3: Comparison of Accuracy on Multinl I Dataset.

Models	SNLI (%)	Matched (%)	Mismatched (%)
Most frequent	34.3	36.5	35.6
CBOW	80.6	64.8	64.5
BILSTM	81.5	66.9	66.9
ESIM[Chen et al , 2016]	86.7	72.4	71.9
SDF-NN (ours)	88.2	77.5	76.5

of reducing the complexity of the model parameters, the optimal model parameters were determined to be 3 functions G and 2 discard gates (dr-gates) .

4.3. Performance Evaluation On The Multinli Corpus. In addition to the more widely used SNLI corpus for training, Stanford has recently released a new corpus, MultiNLI (Multi-Genre Natural Language Inference). The MultiNLI corpus contains 433k pairs of sentences that are annotated with textual implication information. This corpus is modeled after the structure of the SNLI corpus, but differs in that it covers more types of spoken and written language, and thus the dataset has greater diversity and complexity. In particular, the corpus contains samples from more than ten sources, including ten types of corpus from texts, spoken dramas, and so on. The validation and test sets contain samples from all ten types of corpora, but the training set contains only five types. The test data set of the corpus contains two additional types of data. Matched examples and Mismatched examples. Matched examples means that the sample type is present in the test set and also in the training set. Mismatched examples means that the sample type appears in the test set but not in the training set. The performance of the SDF-NN model can be evaluated on both test sample sets, and the relevant parameter settings for the experiments on this corpus are consistent with those on the SNLI corpus.

Table 4.3 shows the evaluation of the different models on this corpus. The first one is a CBOW model based on bag-of-words model and the second one is modeled with a basic bidirectional LS T M. ESIM is the base model without the tree LS T M structure. the SDF-NN model achieves an accuracy of 77.5% on the Matched test set and 76.5% on the Mismatched test set, both outperforming the other models. This indicates that the SDF-NN model has strong learning ability and generalization ability, and is suitable for more complex data sets.

This chapter proposes a memory mechanism for reading comprehension problems. Through the memory correction mechanism, the original document information is fused to correct the attention during the iterative learning process of the network, so that the original information is taken into account in each iteration of the attention, preventing significant attention bias and eventually outputting more accurate answers. Based on this memory mechanism, we designed the memory Gated Attention Reader (mGA) model, an end-to-end neural network-based model, to address the problem of attentional bias that occurs in models based on inference mechanisms. On the CNN dataset, Daily Mail dataset, and CBT data, we show that the prediction accuracy of our model is higher than that of some of the previously proposed models, and validate the effectiveness of the proposed memory correction mechanism. This memory mechanism is also different from previous models such as MemNet, which does not store the combined information of documents and queries in a separate component of the network, but directly introduces the original information into the network iterations repeatedly, solving the problem of information compression and loss caused by the deepening of the network. This approach is also a good analogy to the way the human brain works, where each time the focus of a query is found in an article, it is always based on a global article memory context. Future work will apply this idea of introducing raw information to solve the network information compression problem to other problems and propose a more general memory mechanism.

5. Conclusion. We propose a strategy to combine natural language processing and wireless communication for English classroom translation for university students. In this paper, we propose a semantic discard network and a semantic fusion alignment method for the semantic implication relation problem through the analysis of human thinking process of natural utterance relation determination, and propose an SDF-NN model, an end-to-end neural network model based on these two innovative methods. The SDF-NN model, an end-to-end neural network model, is proposed based on the two innovative methods. The natural utterances often have interference semantics on the final relation judgment, and the semantic discard network can discard the interference information to obtain more accurate local inference results, and then use the semantic fusion alignment method to align the relationship between the local inference results and better fuse these local inference results. The SDF-NN model achieves an accuracy of 88.2% in the public dataset SNLI, which is higher than the single model proposed by other related studies. The SDF-NN model also achieves 77.5% and 76.5% accuracy on the latest dataset, MultiNLI. This demonstrates that the model can learn more complex datasets and has the ability to learn generalization. In general, we provide a new model for the university English translation classroom.Future research can investigate the model's potential for use in other educational domains, such as multidisciplinary language learning and the creation of technological tools for translation. In addition, going over the model's flexibility and potential for generalization in other linguistic and cultural contexts will give readers a more thorough understanding of the model's applicability. Furthermore, it may be worthwhile to suggest avenues for further model modification, such as refining the algorithm to increase translation accuracy and real-time performance and assessing the model's efficacy in practical teaching situations.

Data Availability. The experimental data used to support the findings of this study are available from the corresponding author upon request.

Funding Statement. The Planned Project of Fuyang Social Science Federation: Implement moral cultivation and improve the education in Fuyang City: Long-term mechanism of teacher ethics construction from the perspective of "the Balanced Development of Five-Education" (FSK2023026); The Practical Education Project of Fuyang Normal University: Research on integrated English teaching practice in college, middle school and primary school based on "aesthetic education" (2022SJJYZX09); The Planned Project of Educational Science of Fuyang Education Bureau: Practice and exploration of middle school English teaching mode combining reading and writing from the perspective of "intertextuality" theory (FJK23028).

REFERENCES

[1] ZHU, Qiuyan.*Empowering language learning through IoT and big data: an innovative English translation approach.* Soft Computing, 2023, 27.17: 12725-12740.
[2] Chowdhary, K. *Natural language processing.* Fundamentals of artificial intelligence, 2020, 603-649.https://doi.org/10.1007/978-81-322-3972-7_19

[3] Diao, L.; Hu, P. *Deep learning and multimodal target recognition of complex and ambiguous words in automated English learning system.* Journal of Intelligent & Fuzzy Systems, 40(4), 2021, 7147-7158.https://doi.org/10.3233/JIFS-189543.

[4] Han, R.; Yin, Y. *Head-Driven English Syntactic Translation Model Based on Natural Language Processing.* In 2021 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC) 2021, (pp. 1244-1249). IEEE.https://doi.org/10.1109/IPEC51340.2021.9421111

[5] He, T. *ON TRANSLATION TEACHING THEORY AND TRANSLATION SKILLS IN COLLEGE ENGLISH BASED ON COGNITIVE IMPAIRMENT.* Psychiatria Danubina, 34(suppl 1), 2022, 706-708.

[6] Jingchun Zhou.; Jiaming Sun.; Weishi Zhang.; Zifan Lin. *Multi-view underwater image enhancement method via embedded fusion mechanism.* Engineering Applications of Artificial Intelligence, 2023, 121, 105946https://doi.org/10.1016/j.engappai.2023.105946.

[7] Li, B. *Research on English Translation Based on Recursive Deep Neural Network.* In 2021 3rd International Conference on Artificial Intelligence and Advanced Manufacture, 2021, (pp. 483-487).https://doi.org/10.1145/3495018.3495104

[8] Li, X.; Liu, L.; Tu, Z.; Li, G.; Shi, S.; Meng, M. Q. H. *Attending from foresight: a novel attention mechanism for neural machine translation.* IEEE/ACM Transactions on Audio, Speech, and Language Processing, 29, 2021, 2606-2616.https://doi.org/10.1109/TASLP.2021.3097939

[9] Liu, J.; Tang, B. *GDASC: Assessment of urban land use efficiency in Hebei Province based on data envelopment analysis.* International Journal of Cooperative Information Systems, 2023. https://doi.org/10.1142/S0218843023500132.

[10] Matsui, T.; Suzuki, K.; Ando, K.; Kitai, Y.; Haga, C; Masuhara, N.; Kawakubo, S. *A natural language processing model for supporting sustainable development goals: translating semantics, visualizing nexus, and connecting stakeholders.* Sustainability Science, 17(3), 2022, 969-985.https://doi.org/10.1007/s11625-022-01093-3

[11] Őzcan, F.; Quamar, A.; Sen, J.; Lei, C *Efthymiou, V.: State of the art and open challenges in natural language interfaces to data.* In Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data, 2020, (pp. 2629-2636).https://doi.org/10.1145/3318464.3383128

[12] Radford, A.; Kim, J. W.; Hallacy, C.; Ramesh, A.; Goh, G.; Agarwal, S.; Sutskever, I *Learning transferable visual models from natural language supervision.* In International Conference on Machine Learning, 2021, (pp. 8748-8763). PMLR

[13] Zheng, Y. *Strategies to Improve the Effectiveness of College English Translation Teaching.* Advances in Vocational and Technical Education, 3(2), 2021, 86-91.

[14] Zhu, Qiuyan. *"Empowering language learning through IoT and big data: an innovative English translation approach."* Soft Computing 27.17 (2023): 12725-12740.

[15] Wang, Yuhua. *"Artificial Intelligence technologies in college English translation teaching."* Journal of psycholinguistic research 52.5 (2023): 1525-1544.

[16] ALHALANGY, Abdalilah, et al. *Exploring the impact of AI on the EFL context: A case study of Saudi universities. Alhalangy, AGI, AbdAlgane, M.(2023).* Exploring The Impact Of AI On The EFL Context: A Case Study Of Saudi Universities. Journal of Intercultural Communication, 2023, 23.2: 41-49.

[17] MAHYOOB, Mohammad; AL-GARAADY, Jeehaan; ALBLWI, Abdulaziz. *A proposed framework for human-like language processing of ChatGPT in academic writing.* International Journal of Emerging Technologies in Learning (iJET), 2023, 18.14.

[18] BASKARA, Risang, et al. *Exploring the implications of ChatGPT for language learning in higher education.* Indonesian Journal of English Language Teaching and Applied Linguistics, 2023, 7.2: 343-358.

[19] YOO, Jiseung; KIM, Min Kyeong *Using natural language processing to analyze elementary teachers' mathematical pedagogical content knowledge in online community of practice.* Contemporary Educational Technology, 2023, 15.3: ep438.

[20] YUAN, Qiwei; DAI, Yu; LI, Guangming. *Exploration of English speech translation recognition based on the LSTM RNN algorithm.* Neural Computing and Applications, 2023, 35.36: 24961-24970.

Edited by: Bradha Madhavan
Special issue on: High-performance Computing Algorithms for Material Sciences
Received: Jul 19, 2024
Accepted: Sep 2, 2024