SIN-CHN SCIENTIFIC PRESS PTE.LTD.

Molecular & Cellular Biomechanics 2025, 9(1), 856.
https://doi.org/10.62617/mcb856

Article

Design of an intelligent English learning platform combining biomechanical analysis with biological data analysis and text semantic matching

Hongming Zhu

English Department, School of Foreign Language, Harbin University, Harbin 150086, China; 13351619142@163.com

CITATION

Zhu H. Design of an intelligent English learning platform combining biomechanical analysis with biological data analysis and text semantic matching. Molecular & Cellular Biomechanics. 2025; 22(2): 856.
https://doi.org/10.62617/mcb856

ARTICLE INFO

Received: 21 November 2024
Accepted: 12 December 2024
Available online: 22 January 2025

Abstract: Intelligent classrooms have demonstrated significant promise in enhancing learning efficiency as a result of the quick development of big data and artificial intelligence technologies. This study proposes a text semantic matching model (OM) that combines deep learning and K-means clustering algorithm, aiming to optimize vocabulary. Importantly, it delves into the biomechanical aspects of learning by considering how physical and physiological processes interact with language acquisition. By mimicking the learning mechanism of biological neural networks and further exploring the biomechanical correlates of neural activity during learning, such as the muscle tensions and postural changes associated with cognitive efforts, this model simulates how the brain processes and stores language information. These biomechanical factors can have an impact on concentration and fatigue levels, which in turn affect semantic understanding and memory performance during the learning process. The experimental results indicate that this method not only improves teaching effectiveness, but also provides a solid foundation for future research on intelligent language learning environments, taking into account the biomechanical underpinnings of learning.

Keywords: biological, smart classroom; artificial intelligence; k-means clustering; text semantic matching; biomechanical

1. Introduction

Big data technology and artificial intelligence (AI) have emerged as the main drivers of change in the educational landscape due to the quick advancement of technology. In this context, the construction of intelligent classrooms has become an important means to enhance educational effectiveness. Especially in English education, smart classrooms can not only promote the improvement of students' learning efficiency, but also optimize the teaching process through personalized learning paths and real-time feedback. However, despite the changes brought about by modern educational technology, traditional English teaching still faces many challenges, especially the difficulty of vocabulary learning. Traditional vocabulary teaching methods mostly rely on teachers' explanations and students' memorization. Although this method can improve students' memory ability to a certain extent, it often overlooks their understanding and application of vocabulary in practical contexts. Therefore, how to apply advanced technological means to English vocabulary teaching and enhance students' language cognitive ability has become a key issue in current education reform [1–3].

Language learning, as a complex cognitive process, involves the brain's processing and integration of various information such as speech, vocabulary, grammar, etc. In recent years, neuroscience research has revealed the deep mechanisms of language learning, especially in the process of vocabulary learning and memory, how different regions of the brain work together, and how reinforcement learning promotes long-term memory and contextual application of vocabulary [4–7]. Therefore, deep learning models based on neural networks that simulate the multi-level information processing of the brain in language processing have become important tools for improving language learning effectiveness. This process not only involves understanding vocabulary, but also involves semantic reasoning and association, especially in the challenges of semantic matching and vocabulary classification for English vocabulary. Therefore, how to simulate this cognitive process through intelligent systems and optimize language learning outcomes is an important direction in current research on intelligent education [8,9].

The construction of intelligent classrooms, combined with artificial intelligence and big data technology, can help teachers achieve personalized teaching and provide students with customized learning resources and feedback. Through big data analysis, the system can accurately understand the weak links in students' vocabulary learning process, and adjust teaching content and methods in real time according to students' learning situation. Artificial intelligence, on the other hand, uses technologies such as deep learning and semantic matching to achieve semantic classification and clustering analysis of vocabulary, helping students better understand and apply English vocabulary in different contexts. This intelligent teaching based on cognitive processes can not only improve students' memory effectiveness, but also help them enhance their vocabulary application ability, thereby achieving the goal of language learning [10].

The intelligent classroom design proposed in this article combines deep neural networks (DNN) and K-means clustering algorithm, aiming to optimize the semantic matching process in English vocabulary teaching through advanced semantic matching models and data analysis methods. This model utilizes DNN model for deep semantic matching, and combines *K*-means clustering method to classify and match different lexical semantics. This method can better simulate the human brain's processing and classification of vocabulary during language learning, providing students with a more personalized and accurate learning experience.

2. The method

One of the current research hotspots is the development of an intelligent university English classroom using artificial intelligence and big data. The suggested AI technique matches different kinds of numerical information in English using distinct lexical information by utilizing a deep neural network text semantic matching model (OM). In the meantime, the semantic information of various words is matched using the k-means clustering algorithm.

3. Original model (OM) for text semantic matching in deep neural networks

In order to gather the interaction information of sequence transitions between sentence pairs, a self-supervised learning model is presented based on the current deep text semantic matching model. It then uses a multi-task learning approach to dynamically incorporate this interaction information into the deep text semantic matching task. The structure of this paper is made up of the self-supervised model (SSM) and the original model (OM) (**Figure 1**). To create the link between the two components of the model, the overall framework uses the multi-task learning hard parameter sharing technique.

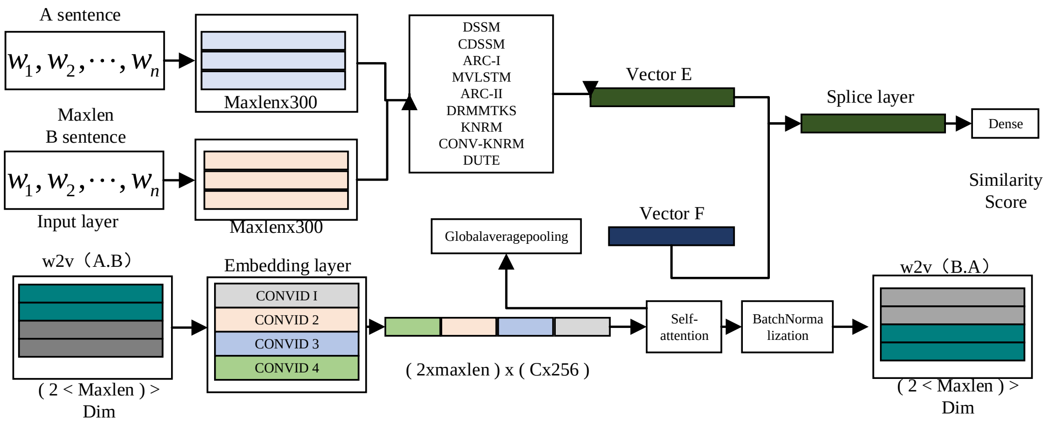


Figure 1. Structure of the model.

Discover how to obtain two phrases' feature interaction vector, Vector_E. This is how Vector_E is computed:

(1) A dataset of n sentence pairs is created by taking two sentence sets, $A = \{a_1, a_2, \dots, a_n\}$ and $B = \{b_1, b_2, \dots, b_n\}$, each containing n sentences. In the i-th sentence pair, indicate the A sentence with $a_i = \{\omega_1^{a_i}, \omega_2^{a_i}, \dots, \omega_m^{a_i}\}$, and the x-th character or word of the sentence a_i with $\omega_x^{a_i} (x \in [1, m])$. (characters for Chinese text, words for English text). In **Figure 1**, m stands for MaxLen, or the sentence sequence's maximum length. Similarly, there is $b_i = \{\omega_1^{b_i}, \omega_2^{b_i}, \dots, \omega_m^{b_i}\}, \omega_x^{b_i} (x \in [1, m])$.

(2) In the experiment, the embedding layer's dimension, or dim, is set at 300. To create the embedded representation, the TSSM embedding layer retrieves the two sentences, ai and bi, of the sentence pair, that is, the matrices $Embed\_a_i \in R^{m \times Dim}$ and $Embed\_b_i \in R^{m \times Dim}$

(3) Enter the embedding representation of the two sentences in sentence pair I into TSSM to get $Vector\_E_i$:

$$Vector\_E_i = TSMM_i(Embed\_a_i, Embed\_b_i) \tag{1}$$

(4) Using the Sigmoid function as the activation function, enter $Vector\_E_i$ into the fully connected layer to obtain the two phrases' Sim_i similarity score:

$$Sim_i = Sigmoid_i(W_o Vector\_E_i + b_o) \tag{2}$$

The parameters that can be learned and changed are W_o and b_o.

Assign the label of the sentence pair to $L = \{y_1, y_2, \dots, y_n\}$, where $y_i (i \in [1, n])$ is the label of the i -th sentence pair. Use binary cross-entropy as the loss function.

$$LOSS_{OM} = -(L \cdot log(Sim_i) + (1 - L) \cdot log(1 - Sim_i)) \tag{3}$$

To address the text semantic matching problem, self-supervised modeling (SSM) extracts the interaction information of sentence pair vector matrices through sequence creation. The pre-training challenge for SSM is to generate sentence pairs with mutual sequences. This is the exact algorithm.

(1) SM input design: two sentences $a_i$and b_i in the sentence pair i are trained using Skip-gram algorithm to generate Word2Vec [11] vector representation, i.e., denotes the length of the sentence sequence and Dim denotes the vector dimension, which is spliced to obtain matrix $W2V\_AB_i \in R^{2m\times \mathrm{Dim}}$.

$$W2V\_AB_i = \begin{bmatrix} W_{a_i} \\ W_{b_i} \end{bmatrix} \tag{4}$$

(2) SSM output design:

$$W2V\_BA_i = \begin{bmatrix} W_{b_i} \\ W_{a_i} \end{bmatrix} \tag{5}$$

SSM takes the sentence pair matrix $W2V\_AB_i \in R^{2m\times dim}$ as the input and labels it as $W2V\_BA_i \in R^{2m\times dim}$.

(3) Convolutional Layer Feature Extraction: a CC-layer one-dimensional convolutional layer (Conv1D) is used to construct a multi-CNN network to extract the n-tuple features of $W2V\_AB_i$ and combine them into a matrix $N_g \in R^{2m\times 256C}$.

$$U_k = Conv1D_k^{k+1}(W2V\_AB_i), k \in [1, C] \tag{6}$$

$$N_g = [U_1, U_2, \dots, U_C] \tag{7}$$

The convolution kernel sizes for the multi-CNN configuration are 2, 3, 4, and 5 for extracting binary, ternary, quaternary, and quintuple features while accounting for multi-character combinations. For Chinese text, (C) is set to 4. In order to extract binary, ternary, and quaternary characteristics from English text, (C) is set to 3, and the convolution kernel sizes are 2, 3, and 4.

(4) Sequence feature extraction and model output: In order to extract the N-element sequence features and ensure that each node's output has complete sequence information, the multilayer convolutional network's output from step 3 is input into the self-attention mechanism. Following the normalized attention mechanism, the temporal fully connected network generates the SSM output, which is then processed by the Softmax activation function.

$$BN = Batch\_Normalization(Self\_Attention(N_g)) \tag{8}$$

$$W2V\_\hat{BA}_i = Soft\,max\,(W_s \cdot BN + b_s) \tag{9}$$

Of these, $W_s$and Bs are learnable and modifiable parameters.

The resemblance between the generated and real vectors is assessed using cosine similarity, which concentrates on the angle between the vectors. However, MAE (Mean Absolute Error) and MSE (Mean Square Error) do not directly show similarity; rather, they concentrate on the difference between the actual and projected values.Cosine similarity serves as the loss function in SSM.

$$Loss_{SSM} = -\cos i\, ne_i(W2V\_BA_i, W2V\_\hat{BA}_i) \tag{10}$$

The basic downstream task (i.e., the original model) receives text interaction data produced by self-supervised learning as part of the multi-task learning framework (OM + SSM) presented in this research. In particular, the vector vector_F is produced by averaging and adding the normalized interaction information (BN) that was taken out of the pooling layer.

$$Vector\_F_i = GlobalAveragePooling_i(BN) \tag{11}$$

The similarity score Sim_Scorei is then obtained by feeding the vector $Vector\_E_i$ of the original model into a fully connected layer that employs the Sigmoid function as an activation function. This is done by splicing the vector $Vector\_E_i$ with the vector $Vector\_F_i$ of the interaction information.

$$Sim\_Score_i = Sigmoid(Wm[Vector\_E_i, Vector\_F_i] + b_m) \tag{12}$$

$$Loss_{ML} = Loss_{OM} + \lambda Loss_{SSM} \tag{13}$$

Table 1 Enumerates the SSM parameters. Each multi-CNN layer has 256 self-attention neurons, with Relu serving as the activation function.

Table 1. Parameters of neural networks.

data set	length of the series	Batch Size	Embedding OM	SSM Input	SSM Output	Multi-CNN
MSRP	35	60	(32,100)	(33,200)	(67,200)	2,3,4,5
CCKS18-T3	39	60	(41,200)	(41,200)	(81,400)	2,3,4,
TCA120	19	60	(20,200)	(20,200)	(41,400)	2,3,4,5
GAIIC21-T3	38	60	(35,200)	(34,100)	(75,200)	2,3,4,
GAIIC21-T3M	29	60	(31,100)	(32,100)	(66,200)	2,3,4,5

To calculate similarity scores for sentence pairings in this work, two models were developed: a decomposition model and a multi-task model that included SSM. In light of this, this research also develops a multi-task SA + SSM model, which weights and adds the loss functions of the SSM and SA decomposition models to determine the multi-task model's anticipated sentence pair similarity. In OM + SSM multitask learning, the weighting is identical to the loss function, where λ takes on a value of 0.5 [12–14].

3.1. The *K*-means clustering approach correlates various lexical semantic data

Finding the constant *K*, or the final number of cluster categories, is the first step in the *K*-means process. Next, using a random selection process, find the centroid by calculating the similarity between each sample and the centroid (in this case, the Euclidean distance); assign the sample points to the class that is most similar to the centroid; recalculate the centroid of each class (i.e., the class center); repeat this process until the centroid stays constant; and lastly, identify the class to which each sample belongs to and the centroid of each class. Because the similarity between each sample and each centroid is calculated each time, the *K*-Means algorithm's convergence speed is relatively slow on large data sets.

The main difference between the classification algorithm and the clustering algorithm is that the latter is a supervised learning method with the aim of determining the result, whilst the former is an unsupervised learning process. The clustering algorithm divides the samples into various groups based on how similar they are to one another. Clustering results will vary depending on the method used to calculate similarity. The Euclidean distance approach is the most often used method for calculating similarity.

After determining each point's distance from every centroid, the nearest centroid is chosen, and the point is then allocated to the appropriate cluster. Each cluster's centroid is recalculated after one iteration, and the nearest centroid is once more determined for every point. After two repetitions, this process is continued until the clusters no longer change [15,16].

First, a class named *k*-means must be defined in order for the k-means algorithm to read and store external data. Then, create a container vector with the data type structure st_point that contains the class's char type ID and 3D coordinates. Next, declare the necessary functions. **Figure 2** displays the flow chart.

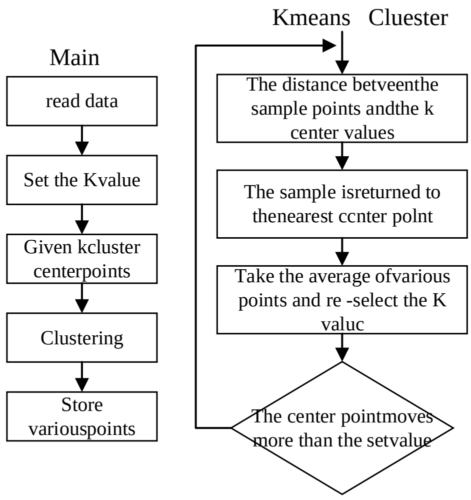


Figure 2. Fundamental program structure and related features.

In *k*-means, the specific provision of different functions of the public function, as shown in **Figure 2**, the function is more complete, easy to follow the expansion of the application. More specific clustering function "clustering" strictly follow the basic principles of *k*-means [17].

The *k*-means algorithm works like this:

1) Select the number of clusters (*k*); only the maximum *k* value should be chosen when passing hyperparameters.

2) Either create *k* centers directly or produce *k* clusters at random and identify the initial cluster centers.

3) The nearest clustering center should be assigned to each point.

4) Check for consistency between the sample points' categories before and after clustering. The algorithm ends if it is consistent; if not, move on to step 5.

5) Determine the sample points' center point for each category, use that as the new center point, and then go to step 2.

3.2. Constructing an Enlightened classroom to teach college English terminology

3.2.1. Pre-class warm-up

Pre-lesson vocabulary warm-ups are designed to help students master the basics of the words they are learning, such as pronunciation, word sense identification, synonym analysis, lexical properties, and grammatical functions, thus laying the groundwork for extended vocabulary explanations and practical use in the classroom. The pre-class warm-up is usually divided into three parts:

1) Information pushing: Teachers can push videos, audios, pictures, memorization methods and word explanations to students through Dingding platform, WeChat groups, emails and other online platforms to help students clarify the learning objectives and key points of the words.

2) Vocabulary tests: Teachers design tests based on teaching objectives to assess students' mastery of vocabulary and set targeted themes. For example, students can be asked to upload audio files and take dictation, or they can be tested on their understanding of word classes, grammar and word meanings through multiple choice fill-in-the-blank questions. Through these tests, students will be able to get a clear picture of their strengths and weaknesses in word mastery and be prepared for classroom learning.

Implications for Teachers: Teachers can use online assessments to learn more about their students' vocabulary mastery, gather firsthand information, and compile a list of the words that pupils have mastered and those that still require improvement. Teachers can provide more exercises and explanations for language that students are struggling with in the classroom. To maximize and standardize vocabulary instruction, teachers should simultaneously keep updating materials for vocabulary explanations and investigate more efficient teaching strategies.

3.2.2. Cooperation in the classroom

In order to assist students feel and experience language at a higher level and gain a deeper understanding of its nature, the class largely focuses on collaboration and interaction between teachers and students, with a particular emphasis on vocabulary output. The pre-class warm-up is expanded upon, and pupils' vocabulary is significantly improved. Level 2 testing, student contact, and teacher-student interaction are all elements of the classroom.

3.2.3. After-class comments

After the session, the teacher records the students' learning and assists them in reviewing, which is known as post-lesson feedback. Teachers can assign pupils to one of three grades—excellent, average, or poor—based on the outcomes of their second vocabulary test. Teachers at different grade levels will use different ways, degrees and contents of tracking depending on the students.

For top students, they have mastered and flexibly used the vocabulary they have learned, thoroughly internalizing it. Intermediate students, on the other hand, have mainly mastered the pronunciation, lexical properties, and meaning of vocabulary, but have difficulty with practical applications, such as sentence formation, conversation, or writing. Lower-advanced students have greater problems with basic vocabulary skills, and teachers should pay more attention to and help them, especially in the reinforcement of basic knowledge.

4. Search and examination

Our method combines neural networks with dictionary classification models, which includes the following steps:

1) auxiliary dictionary construction: constructing multiple dictionaries, including entity linking dictionary, participle dictionary, Participle and word frequency computation using a dictionary of words and attributes.

2) Identification of entities and attribute values: Determine the values of entities and attributes. Because the problem's attribute values are less standardized, they can contain lengthy word sequences or not be able to directly match knowledge base items. Some of the entities in the split-word dictionary can be disregarded.

3) entity linking and filtering: compute features for each entity and perform linking and filtering.

4) Generate Candidate Query Paths: perform text matching to generate candidate query paths.

5) entity splicing and answer retrieval: complete entity splicing and retrieve answers.

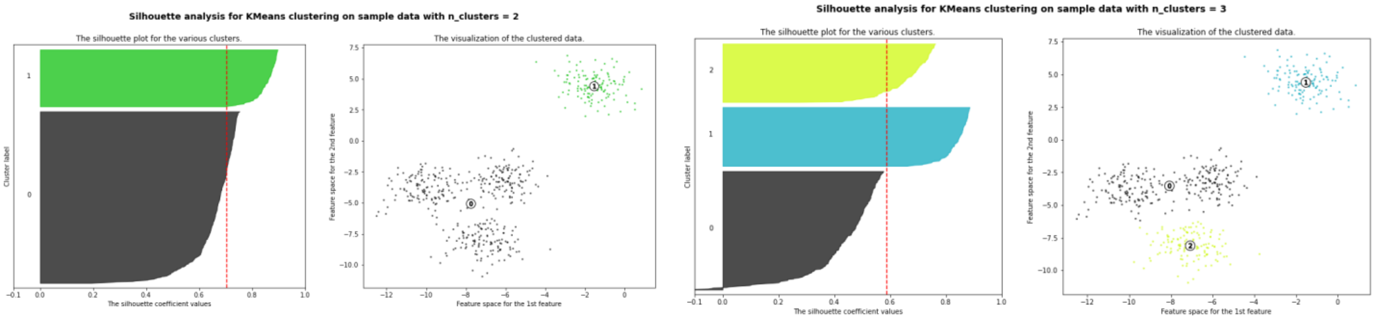
Table 2 presents the data statistics.

Table 2. data set statistics.

TYPE OF QUESTION	Training set	Validation set	Test set
One entity, one relationship	1145	465	488
Numerous relationships with a same entity	675	154	162
multi-entity	355	132	120

4.1. Clustering of data

The experiment moves onto the clustering phase following the completion of word vector transformation and data pretreatment. The silhouette coefficient typically has a range of [−1, 1], and the bigger its value, the more compact the clusters are inside and the farther apart they are from one another. Furthermore, the persuasiveness and representativeness of the clustering results will be impacted by an excessive or insufficient distribution of data in a single cluster. Consequently, the silhouette coefficient and the scatter plot's data distribution must be taken into account when calculating the K value. In other words, if the distribution of the clustered data is more uniform, the corresponding K value is more appropriate based on the maximum value of the silhouette coefficient. This study conducted several sets of controlled experiments to determine the best clustering distribution results; the precise outcomes are displayed in **Figure 3**.



(a) **(b)**

Figure 3. The GMM scatter plots and contour coefficient values with various clustering numbers. **(a)** The silhouette coefficient value is 0.008 and there are three clusters; **(b)** there are four clusters and a silhouette coefficient of −0.035.

It is evident from examining the experimental data that the silhouette coefficient is comparatively high when there are two or three clusters, but it progressively drops as the number of clusters rises. The scatter plot demonstrates that the distribution of three clusters is more uniform when $K = 3$, whereas the distribution of two clusters is unequal when $K = 2$. Thus, based on a thorough comparison of the experimental findings, three clusters are the ideal number for the data gathered for this paper. At this point, the center point is utilized as the clustering center, and the parameter K is set to three for the final K-means clustering. **Figure 4** displays the sample data's clustering impact.

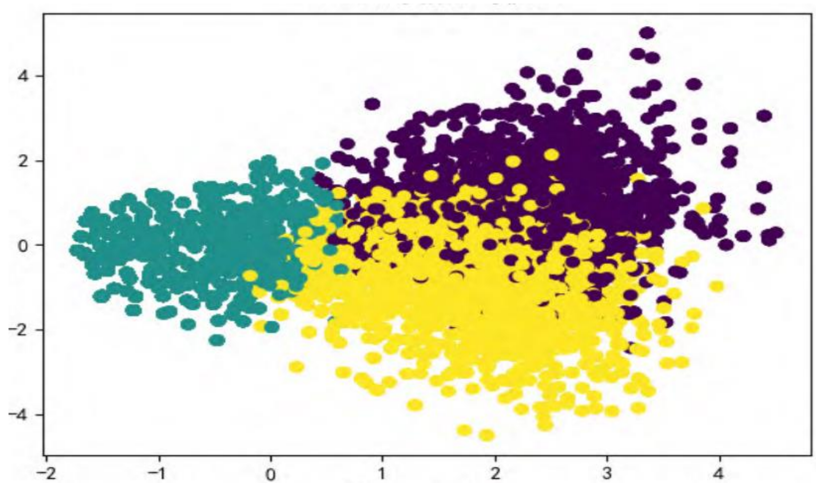


Figure 4. K-means scatter diagram for clustering.

The number of clusters is set at three, and three conventional clustering algorithms—the Birch, Hierarchical, and DBSCAN algorithms—are chosen at random while maintaining all other parameters constant. Both graph and evaluation index viewpoints are used to analyze the experiment findings. **Table 3** demonstrates that the G-K-means algorithm has the highest profile coefficient, the greatest CH index value, and the smallest DB index value. This suggests that the algorithm's clustering impact is superior to that of the other algorithms and more consistent with the "optimal clustering quality" approach. The Birch, Hierarchical Clustering, and DBSCAN algorithms, on the other hand, have poor distribution and poor clustering effects. In particular, the DBSCAN method has trouble dividing the data set into manageable clusters, which makes it clearly unsuitable for clustering the text of online Q&A community questions.

Table 3. Comparative analysis of several algorithms' evaluation indices.

TYPE OF ALGORITHM	DB Index	CH Index	THE SILHOUETTE FACTOR
G-*K* means	4.5567	1668.25	0.2561
birch	4.785	1326.82	0.2672
DBSCAN	462.54	37.452	0.3668
Clustering in Hierarchy	6.325	1324.51	0.2145

4.2. Deep learning in conjunction with the original model (Om) for text semantic matching

The experimental findings in **Table 4** are used to discuss the two research issues raised in this work. Accuracy, AUC, and F1 score range from 0 to 1. First, RQ1 is discussed. These findings show that the addition of SSM enhances the performance of the representation-based models on the five datasets and that the interaction information gleaned by self-supervised learning successfully makes up for these models' drawbacks.

Table 4. Model performance comparison (%).

	MODEL	MSRP	CCKS18-T3	TCAI20	GAIIC21-T3	GAIIC21-T3M
OM SSM multi-task model and representation-based OM model	ARC-I	75.26	68.26	75.49	84.62	90.62
	ARC-I + SSM	76.52	71.65	76.58	85.62	93.65
	DSSM	82.14	74.26	65.88	75.62	80.78
	DSSM + SSM	81.16	75.66	84.62	88.98	90.99
	CDSSM	79.86	71.46	72.66	78.91	68.52
	CDSSM + SSM	81.36	77.68	85.02	88.86	90.62
OM SSM multitasking model, hybrid OM model, and interaction-based OM model	ARC-II	76.52	71.36	74.26	81.45	84.66
	ARC-II + SSM	78.69	70.36	77.68	86.36	91.68
	DRMMTKS	77.98	74.26	87.65	66.24	68.32
	DRMMTKS + SSM	80.89	76.52	87.89	86.26	90.26
	K-NRM	77.98	73.65	72.58	66.26	62.32
	K-NRM + SSM	78.58	72.36	81.26	88.36	92.35
	CONV-KNRM	76.36	77.25	76.36	82.36	80.24
	CONVKNRM + SSM	81.26	75.68	85.62	85.68	89.96
	MVLSTM	75.26	74.26	78.86	80.26	82.53
	MVLSTM + SSM	78.63	72.36	82.14	88.25	92.63
	DUET	75.68	75.62	84.63	80.35	4.68
	DUET + SSM	77.69	75.62	84.63	87.41	92.36
SA + SSM multitasking and the decomposition model	Self-attention(SA)	81.36	74.62	80.26	89.36	92.63
	SA + SSM	80.25	74.62	83.20	87.52	93.23

The performance improvement for each dataset following the combination of the self-supervised models is shown in **Table 5**. With an average improvement of 2.44%, the nine models' improvement for the MSRP dataset is more constrained. This dataset's sentences originate from various news websites and news articles, which effectively reduces the potential semantic similarity between sentences. This reduces topic complexity and sentence commonality, indicating that SSM is not robust enough to generate sentence pairs that address different topics. The model's improvement with the addition of self-supervised learning is more evident for the other three datasets. GAIIC21-T3(M) focuses on brief text matching for AI-assisted interactions, while TCAI20 focuses on judging comparable sentences associated with the New Crown outbreak. SSM is able to extract high-quality interaction information from these datasets, which contain more comparable sentence topics.

Table 5. Data-set-based self-supervised model enhancement (%).

Dataset	**ARC-ISSM**	**DSSM + SSM**	**CDSSM + SSM**	**ARC-IISSM**	**DRMMTKS + SSM**	**KNRM + SSM**	**CONV + SSM**	**MVL + SSM**	**DUET + SSM**	**AVG**
MSRP	2.9	0.4	1.2	1.8	2.8	2.2	3.2	5.2	2.5	2.54
CCK-T3	2.8	5.2	8.6	0.2	2.6	4.6	2.2	3.2	2.5	3.45
TCA120	2.6	21.5	18.8	1.7	1.2	12.5	17.8	4.9	2.8	9.02
GAIIC21-T3	3.8	15.8	11.5	5.5	27.4	32.6	4.5	8.7	8.6	13.65
GAIIC21-T3M	1.5	12.6	30.8	7.4	35.9	45.6	11.6	12.4	9.5	17.88

4.3. English corpus information matching

Ablation experiments are conducted on the test set for five different feature types for the entity linking link, and the recall rates of keeping varying numbers of candidate entities are documented. The recall rate of all question-annotated entities while keeping the first n candidate entities is represented by Recall@n/% in **Table 6**, which displays the experimental results.

Table 6. The entity link's outcome on the test set.

feature	**reserved quantity**	**Recall@n/%**
Entity feature plus question entity	All (avg 12.6)	94.6
Entity feature plus question entity	2	82.4
Entity feature plus question entity	4	93.6
Entity feature plus question entity	6	94.8
Entity feature plus question entity	8	95.6
Only solid characteristics	4	87.4
Features of Question-Only Entities	4	72.6

F-values were calculated on the test set for various counterexample counts and retrieval strategies. Three scenarios' performances are compared in this paper: After text matching, (1) the query path with the highest similarity is directly chosen; (2) a bridging technique is used to generate potential query paths for each question in the multi-entity case; and (3) the first three paths of the text matching are re-matched with the multi-entity paths by overlapping words and choosing the path that is most similar to the final query path.

The advantages of the model include (1) with the help of pre-training model and knowledge base segmentation technique, it significantly improves the recognition accuracy of question subject terms; (2) through text matching technique, it matches the query paths of questions and knowledge base entities, thus avoiding the problem of unregistered relations; (3) through entity splicing method, it can effectively deal with multi-entity and multi-relationship problems. The drawbacks, on the other hand, include (1) the model relies on the features of the question entities and the knowledge base entities, and thus has high requirements for machine learning-based entity linking techniques; (2) the creation of more potential query pathways, which impacts the model's effectiveness. According to the authors, entity type and entity number information can be added to the problem to increase the accuracy of multi-entity and multi-relationship problems, and deep learning techniques can be used to link entities in the future to decrease feature dependency and improve accuracy.

5. Conclusion

The integration of biological principles with modern AI technology in smart classrooms offers a novel approach to language learning. From a biological perspective, language acquisition is a complex cognitive process involving various brain regions responsible for memory and semantic processing. Traditional vocabulary learning methods are limited in optimizing these brain functions. Smart classrooms, utilizing deep learning and big data, simulate brain-like processing of information, enhancing vocabulary learning through contextualized semantic matching and clustering algorithms.

This technology mirrors the brain's associative memory and cognitive functions, creating personalized learning paths that align with individual neural response patterns. By optimizing cognitive processes such as memory retention and semantic understanding, smart classrooms foster more effective learning. Ultimately, this model not only enhances language acquisition but also contributes to understanding the biological mechanisms of learning, marking a significant advancement in educational practices.

Ethical approval: Not applicable.

Conflict of interest: The author declares no conflict of interest.

References

1. Ali, R.; Muhammad Hameed, S.; Sungyoung, Lee.: Rough set-based approaches for discretization: a compact review. Artificial Intelligence Review 44, no. 2, 2015, 235-263. https://doi.org/10.1007/s10462-014-9426-2
2. Alsubari, S. N.; Deshmukh, S. N.; Alqarni, A. A.; Alsharif, N. H. T.: et al. Data Analytics for the Identification of Fake Reviews Using Supervised Learning. CMC-Computers, Materials & Continua, 70(2), 2022, 3189–3204. https://doi.org/10.32604/cmc.2022.019625
3. Chen, H. C.; Chen, Z. Y.; Huang, S. Y.: et al. Relation extraction in knowledge base question answering: From general-domain to the catering industry, //International Conference on HCI in Business, Government, and Organizations. Springer, Cham, 2018, 26-41. https://doi.org/10.1007/978-3-319-91716-0_3

4. Dai, Z.; Li, L.; Xu, W.: Cfo: Conditional focused neural question answering with large-scale knowledge bases, arXiv preprint arXiv:1606.01994, 2016. https://doi.org/10.18653/v1/P16-1076
5. Devlin, J.; Chang, M. W.; Lee, K.: et al. Bert: pre-training of deep bidirectional transformers for language understanding // Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Minneapolis, 2019, 4171–4186.
6. Graves, A.: Long short-term memory, Supervised sequence labelling with recurrent neural networks, 2012, 37-45. https://doi.org/10.1007/978-3-642-24797-2_4
7. Hu, S.; Zou, L.; Zhang, X.: A state-transition framework to answer complex questions over knowledge base, //Proceedings of the 2018 conference on empirical methods in natural language processing. 2018, 2098-2108. https://doi.org/10.18653/v1/D18-1234
8. Li, B.: Application of flipped Classroom Teaching Model based on micro-lesson in High School English Grammar Teaching, Campus English, (38), 2020, 162-163.
9. Liu, Q.; Liu, C.; Wang, Y.: etc. Integrating external dictionary knowledge in conference scenarios The field of personalized machine translation method , Journal of Chinese Informatics, 33(10), 2019, 31-37.
10. Lv, M.: Research on The Teaching Mode of Teacher-student Collaborative Development in High School English Teaching, New Curriculum Research, (35), 2020, 89-90.
11. Ning, C.; Xu, J.; Gao, H.: et al. Sports Information Needs in Chinese Online Q&A Community: Topic Mining Based on BERT, Applied Sciences, 12(9), 2022, 4784. https://doi.org/10.3390/app12094784
12. Palanisamy, S.; Thangaraju, B.; Khalaf, O.I.; Alotaibi, Y.; Alghamdi, S.; Alassery, F.: A Novel Approach of Design and Analysis of a Hexagonal Fractal Antenna Array (HFAA) for Next-Generation Wireless Communication. Energies 2021, 14, 6204. https://doi.org/10.3390/en14196204
13. Rezgui, Y.; Boddy, S.; Wetherill, M.: et al. Past, present and future of information and knowledge sharing in the construction industry: Towards semantic service-based e-construction, Computer-Aided Design, 43(5), 2011, 502-515. https://doi.org/10.1016/j.cad.2009.06.005
14. Turnbull, M.: There is a role for the L1 in second and foreign language teaching, The Canadian Modern Language Review, 57(4), 2001, 531-540. https://doi.org/10.3138/cmlr.57.4.531
15. Wang, J.; Wang, L.; Xu, J.: et al. Information needs mining of COVID-19 in Chinese online health communities, Big Data Research, 24, 2021, 100193. https://doi.org/10.1016/j.bdr.2021.100193
16. Wang, Y.; Zhang, R.; Xu, C.: et al. The APVA-TURBO approach to question answering in knowledge base, //Proceedings of the 27th International Conference on Computational Linguistics. 2018, 1998-2009.
17. Wu, B.: The Reform of English Teaching Model in High School under the Background of "Internet +" , Anhui Education and Research, (04), 2020, 80-81.