

¹Zenghui Du

Optimization Analysis of AI Intelligent English Teaching Strategies Based on Computer Virtual Reality Technology



Abstract: - The Internet of Things and other technology breakthroughs have a big impact on how English is taught (IoT). The investigation of the actual English teaching process and the identification of student features come first. In order to use the interpolation technique with the student image-based feature detection method, we investigate the IoT reform. As a result, we find a clever algorithm for IoT that recognises student features more effectively. Artificial intelligence (AI) is the design technique used to create the intelligence algorithm for pupil feature recognition. A comprehensive multifunctional human-computer interaction system utilizes a variety of input and output streams. In addition to the standard computer keyboard, cursor clicking, and screen touching, the most recent speech and facial recognition technology can be employed for data input. Students learned to orally interact with the robot and act as a guide to various destinations. The Multimodal Interaction System for English Education (MMIEE) is an investigation into the use of network and artificial intelligence (AI) in teaching. Recurrent neural network (Rein RNN)-based Reinforced learning is utilized for the perpetual evaluation model to categorize the questions posed by teachers in terms of their content and kind, and to conduct experimental investigation. The results show that the proposed Rein_RNN model achieves 98.6% of accuracy in 4.3sec of mean evaluation time.

Keywords: English education, neural network, multimodal teaching, artificial intelligence, human interaction, online course.

I. INTRODUCTION

The first setting in which kids learn and develop in a school setting is the classroom[1]. Yet, it has been discovered through years of teaching experience and experience summarising that the amount of classroom engagement will also have a significant impact on information acquisition. Due to the uniqueness of the language and the impact of noncommunication, the English classroom in particular has caused significant disruption to the interaction of the English teaching classroom[2]. On the one hand, the entire process and effectiveness of instruction are diminished by students' inconsistent acceptance of multimodal learning. Yet, despite the addition of new teaching methods, the difficulty of the English topic itself has not diminished, leaving the students' initial issues unresolved[3]. So that students can experience immersive learning, this paper's study and analysis centre on how to utilise all of the qualities of the English subject[4]. In order to create learning methods that are appropriate for each type of learner through prediction and learning methods, this research also mixes artificial neural networks[5]. The artificial neural network-based multimodal teaching interaction model can make the English classroom interesting and fun. In this school environment, learners can get rid of the conventional passive studying state inside one fell swoop, thus transitioning into a positive mental attitude[6]. This can help pupils become more fluent in all aspects of English while also developing their whole personality and traits. An MI model utilizes and organizes multiple modalities to enable for input data to be provided to an agent. The MI system can also produce data in a variety of modalities. A person or a computer can be the representative that communicates with the MI system. Advanced technologies have enhanced the prospect of creating novel MI systems that can connect the physical and virtual worlds[7]. Virtual information modelling in real-time with a context-sensitive concentrate on the data of the representative has become accessible due to the advancement of networks and computer power, which enables a wide range of devices to access the Internet with enough level of performance. The technology that achieves this goal is known as the Internet of Things (IoT), which connects the Online to information and communication techniques (ICT) such sensors that are placed on real-world items [8]. It is anticipated that more and more physical objects will be fitted with ICT and transformed into IoT [9], opening up chances to use interaction modes in education in novel ways. In view of this the contribution of this work are as follows:

¹Nanjing Vocational University of Industry Technology, Nanjing, Jiangsu, China, 210023

*Correspondence should be addressed to Zenghui Du; 2001100232@niit.edu.cn

- To merge artificial neural networks with multimodality by developing an interactive English teaching approach that relies on the pair, which really is crucial for the unification of instructional content and the modernization of instructional strategies.
- To use Recurrent Neural Network (Rein RNN)-based Reinforcement learning to categorize the English teaching-based data collected from 8,957 students, with a significant amount of 37,987 discussion groups.
- The robot successfully interacted with a human after training with datasets created as a series of temporal flows of human-robot interaction. It did this by switching between the acknowledgment, generation, and waiting phases on its own and by making use of the systematically acquired relationships in the right contexts using only forward calculation of the Rein_RNN.
- We tested our approach by having a robot produce suitable behaviour in response to language instructions from a person. After learning, the network actually created an attractor structure that, in its internal dynamics, represented both the temporal rhythm of the task and links between language and behaviour.
- When used effectively, computers in the classroom can be advantageous to both teachers and students. These might make it easier for students to assimilate the knowledge they are learning.

The structure of the current document is as follows: A relevant collection of research for English education using neural networks is provided in section 2 of the presentation with table. In section 3 suggested Recurrent neural network (Rein RNN)-based Reinforced learning classifier is given. In part 4, the performance of the suggested model is shown along with a benchmark methods . Fifth section presents the general conclusion for the suggested method.

II. RELATED WORKS

We provide a comprehensive analysis of multifunctional human-computer interaction (HCI) with the major aims of demonstrating that various technological platforms are applied in various topic areas, or domains. Currently, a number of domain-specific polls and studies have been released.

According to [10], fuzzy neural networks play a significant role in intelligent information processing and can successfully handle nonlinear and fuzzy situations. It additionally has the potential to comprehend how humans function and respond to diverse circumstances. In accordance with a thorough overview of the academic English class in the Mobile learning platform amended version (such as the organisation of basic principles of recent adaptations), a skin colour allocation prototype for colour image pre - processing is established in [11] using a regulation testing regime. This model serves as a warranty and assistance for the education system and student achievement review process. Within [12], a Mel frequency cepstrum coefficient window - based function and the sniffing method serve to construct an interactive computer English voice recognition tool. The system has the capability of swiftly and correctly recognising English voice data, and it uses reference databases for assessment and excellent knowledge databases for language fault rectification, considerably increasing recognition efficiency. Though the investigation and evaluation of college students' English learning versatility endorsed by AI, [13] Long Short Term Memory based analysis aids in delineating its impact the different key variables on gaining knowledge flexibility and the correlation among learning flexibility and various factors, as well as recommending methods for enhancing students' educational flexibility. [14] suggests a collaborative, artificial neural network-based technique for educating multimodal English. It intends to investigate ways to employ artificial neural networks' learner autonomy to speed up the merging of many channels and, at the same time, give recommendations for various instructional interaction modes. Convolution neural networks are designed and implemented with a command prompt data processing platform in [15] to gather student perceiving data about devices in their learning environment as well as statistics on how they interact with touchscreens based on a virtual simulation experiment. The command line sensor data is then processed for learning behaviour using the enhanced neural networks. [16] proposes a model for assessing the efficacy of learning using the neural network's RBF algorithm. Using knowledge integration and improvement of the RBF neural network decisions methodology, the university English teaching performance assessment system has been investigated using the RBF regularisation network method, RBF neural network decision algorithm, and empirical research method. An innovative flipped learning technique for college English was proposed within [17] paper and therefore is based on big data and neural networks with deep layers. As the study participants, 240 second-year English majors from 2 classes overall have been selected to participate in the research. A deep learning-based efficient network education system model for speech processing and face emotion recognition was proposed in [18]. Also, this

research utilised facial expression detection as its primary technology and emotion computation as its conceptual framework to assess and comprehend the psychological response of online learners by capturing and recognising their facial expressions.

Table-1 evaluation of existing methods

Author/year	Method	Advantage	disadvantage
Gu et al., (2023)	Fuzzy neural network	Most powerful model	More cost
Li et al.,(2023)	demand-based testing method with artificial intelligence	Interfacing concept adds more sustainability	Expensive and slow
Deng et al., (2023)	using nose algorithm and Mel frequency cepstrum coefficient windowing function	Effective for word disambiguating	More complex
Bai et al., (2022)	LSTM	More precise results	Number of iterations is more
Hua et al., (2022)	artificial neural networks	More accuracy	More cost
He et al., (2022)	Convolution Neural Network	Needs statistical factor	More computation cost
Chen et al., (2021)	Radial Bias Function Neural networks	Access to multiple layers	Overfitting problem arises
Chang et al., (2021)	Deep neural network	More precision	Number of layers are more
Hao et al., (2021)	Deep learning speech enhancement algorithm	Less complex in nature	Less speed

The study of the literature shows that there are still lot of areas in the reforming colleges and universities that need to be enhanced. Its educational paradigm still has to be thoroughly adjusted to fulfil the requirements of social and economic development, and both the teaching environment and the underlying system are inapplicable. Additionally, the quality of the instruction needs to be improved. In this research, neural network technology will be introduced to the evaluation system, and studies will be set up to confirm its impact. The goal is to implement a multifunctional human-computer interaction (HCI) evaluation process at colleges and universities to support teachers' instruction and enhance the standard of education.

III. PROPOSED METHODOLOGY

Initially, the instructor uploads the course materials and quizzes, which include input modalities like bio signals, sensors, visual signals, audio signals, and tactile objects.

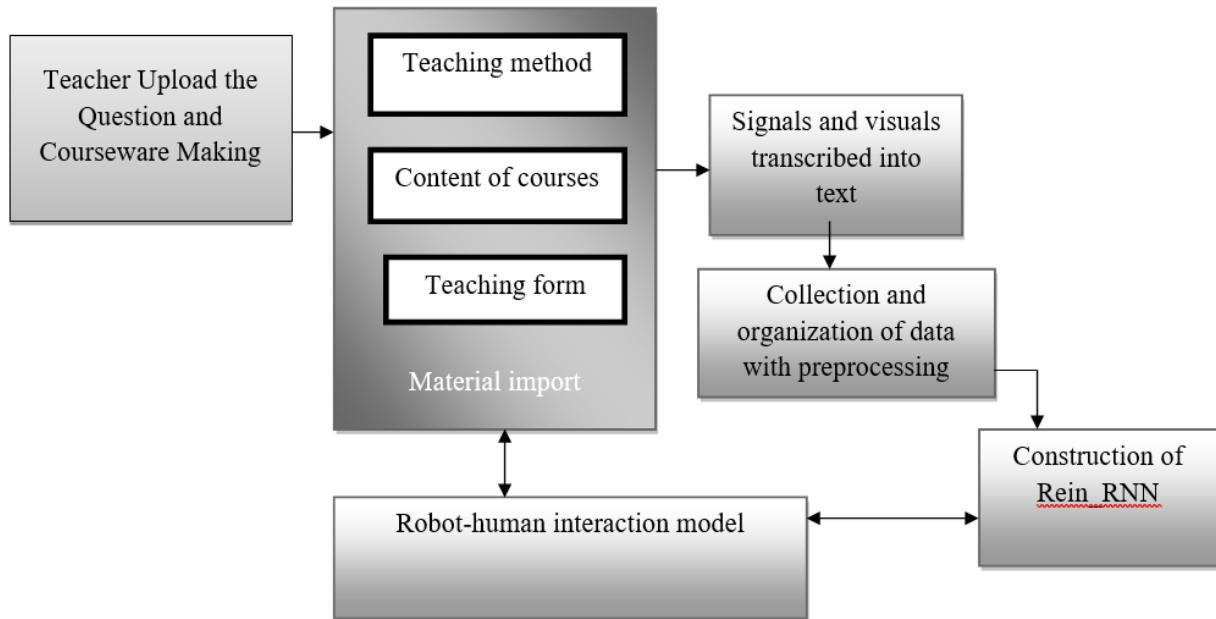


Figure-1 conceptual framework for artificial intelligence based teaching model

The inputs from the various modes are converted to text. As a result, processing and storage are simple. This document with material intake can be organized and stored in a database after data preprocessing. Finally, by implementing a Recurrent Neural Network based on Reinforcement Learning to classify the type of query, as illustrated in figure 1.

Categorization of questions with various input modalities

AI technology is employed in classroom instruction. This encourages the construction of smart classrooms. Additionally, new demands as well as difficulties for AI are also provided by education and teaching. In English classes, questions are treated as research samples. The questioning corpus used in the class is segmented into categories based on the questions' types and contents, as well as the properties of the corpus used in regular classroom instruction and the students' cognitive abilities. The content that is being queried is broken down into three categories: "0" stands for basic knowledge points, "1" for topical data, and "2" for self-management. Types of information points use queries relating to specific knowledge points to look into how well students understand fundamental concepts and can reason and analyses information. Depending on the details of the subject, the teacher decides what kind of material to provide. The question types are denoted with a "0" for basic memory, a "1" for a prompt, a "2" for an analytical, a "3" for an applicability, and a "5" for an assessment. A memorising type question is one that the instructor poses to promote learning in remembering what they have studied. The teacher will use the prompt type to direct students' analysis of a particular subject during class. Teachers often define analytical concerns in terms of particular phenomena. Applied inquiry is when a teacher challenges a class to work together to resolve a different challenge using previously learned concepts and skills. Teacher-asked evaluation-style questions are ones that list the pupils' knowledge. A number of properties, such as bio signal (B), sensors (S), visual signal (VS), sound (SO), and tangible object (TO), are associated to input modalities (TO). We evaluate the impacts of features like time-varying fields in space, together with the feasible approaches are $B(x, y, z, t), S(x, y, z, t), SO(x, y, z, t), VS(x, y, z, t), TO(x, y, z, t)$ where $x \in R, y \in R, z \in R, t \in R$. The notations $u_B, u_S, u_{SO}, u_{VS}, u_{TO}$ are utilized to indicates the intelligent controller's control parameters, then the following system model can be developed::

$$\frac{\partial B}{\partial t} = F_B(B, S, VS, SO, TO, u_B) \tag{1}$$

$$\frac{\partial S}{\partial t} = F_S(B, S, VS, SO, TO, u_S) \tag{2}$$

$$\frac{\partial SO}{\partial t} = F_{SO}(B, S, VS, u_{SO}) \tag{3}$$

$$\frac{\partial VS}{\partial t} = F_{VS}(B, S, VS, u_{VS}) \tag{4}$$

$$\frac{\partial TO}{\partial t} = F_{TO}(u_{TO}) \tag{5}$$

where $F_B, F_S, F_{SO}, F_{VS}, F_{TO}$ are undefined functional. These uncertain functional have precise expressions in an ideal situation. First, L is not affected by other factors, therefore the expression can be made simpler, let $\mu = (B, S, VS, SO, TO)$, $u_\mu = (u_B, u_S, u_{VS}, u_{SO}, u_{TO})$ and by discretizing each variables, it obtains the state sequence $\langle \mu^{(N)}, L^{(N)} \rangle$ and control sequence $\langle u_\mu^{(n)}, u_L^{(n)} \rangle$.

Reinforced learning based Recurrent Neural Network (Rein_RNN)

Assuming for the context of this research that the organized data input feature type is m and that n neurons are added to the network as depicted in figure-2, the data set of the neural network is indicated as shows below:

$$x(t) = (x_1(t), x_2(t), \dots, x_m(t))^T \tag{6}$$

Following is a list of the data type's output data following calculation using the computational formula:

$$y(t) = (y_1(t), y_2(t), \dots, y_m(t))^T \tag{7}$$

The connections between various neurons within the infinite-depth neural net are then written as k for the neuron and as, where $w = [w_u, w_1]$, w_u denote the weights of any interconnected neurons in the network and w_1 indicates the weights of the connections between neurons inside the network and the neurons outside of the network, and then the information on the query distribution that was acquired by the neural network at any given time, $t + 1$, is reported as follows:

$$s(t + 1) = w_u y(t) + w_1 x(t) \tag{8}$$

Equation (8) becomes the following when seen in the time dimension:

$$s(t + 1) = (s_1(t + 1), s_2(t + 2), \dots, s_n(t + 1))^T \tag{9}$$

The outcome of the infinite depth neural net model's dataframe at any given time, following computation by the network, can be recorded as follows:

$$y_k(t + 1) = f_s(s_k(t + 1)) \tag{10}$$

where f_k is activation function of the neuron inside the infinite-depth neural networks. The infinite depth neural network method has the ability to be made larger. The optimization stage is often used to choose a specific number of students to grasp with knowledge base at probability $s = (s_1, s_2, \dots, s_n)$ sampling (along with the categorization level of every knowledge value), where s represent for the relevant classroom factors and is expressed in

$$s = \sqrt{s_1^2 + s_2^2 + \dots + s_n^2} \tag{11}$$

Give each person a knowledge location at random from the study room in order to identify the representative. It is challenging to enable multi-functional education experts to hold a knowledge position at the exact same time. In the education process, replace k . (for every learning representative). The examiners use the discovered knowledge assessments from different levels of understanding for learning, and this education will assist them to understand the pupils. An audience with better location knowledge may recognise the recruiting process and comprehend the fundamental ideas about selection. The following equation represents this alternative method:

$$k = \left[\frac{s_1}{s}, \frac{s_2}{s} \right] \tag{12}$$

The person acquires knowledge of the attribute search idea, switches to new information, and stores the new knowledge in the system $\cos(\theta) = \left(\frac{s_1}{s}\right)$ and $\cos(\alpha) = \left(\frac{s_2}{s}\right)$. From Equation (13), it is clear that decreasing the number of knowledge points (k) in the computer-human interactive class for the minimum achievable level helps students focus more throughout class.

$$k = (\cos(\theta), \cos(\alpha)) \tag{13}$$

Data from the internet survey are used to analyses how the computer-human interaction classroom education technique's design, integration, and scoring system affect each student's ability to study at their own pace. Employing the course materials to prepare the analysis reveals this learning potential. It has advantages for enhancing objectivity, including, among many other things, measuring learning effectiveness using the parameter sp analysed. It is a typical method for assessing a student's capacity for independent learning. Instead of specific academic strategies like in Equation (14). One can compute the value of as:

$$s.p = sp \cos(\theta) \tag{14}$$

This is a frequent method for keeping track of general learning abilities instead of specific learning actions. The present degree of flipped classroom educating techniques is determined by executing by the Equation (15) from $s.p$ teaching system. The purpose $(s_1p_1 + \frac{s_2p_2}{sp})$ is to instruct learners in a more organized and straightforward manner on how to incorporate established educational values and conduct that is suitable for particular spontaneous thinking.

$$s.p = sp \frac{s_1p_1 + s_2p_2}{sp} = s_1p_1 + s_2p_2 \tag{15}$$

Equation (16) asserts that the present level of flipped classroom educating design functions similarly in terms of incorporating students with the educational system. The RNN layer is comprised of four independent RNNs that are used to create state output by applying them to the coupling matrix A. Let the educational observing value sequence state be $\langle \mu^{(n)} \rangle$, The loss function can therefore be defined as follows:

$$\begin{aligned} loss_{\mu}(A, N_p, \langle u_{\mu}^{(k)} \rangle m^{n+1} \langle u_L^{(k)} \rangle m^{n+1}) \\ = \frac{1}{m} \sum_{n=no+1}^m \|\mu^{(n)} - \mu^{1'(n)}\| \end{aligned} \tag{16}$$

where N_p is neutral network parameter. To minimize the target, add up the loss functions of various approaches.

Human machine interaction model

We utilise a miniature humanoid whose body only roughly corresponds to the upper part of the human body for the learning experiment. An I/O neural unit is allocated to each word, and a sentence is represented as a list of words. The head mount camera of the robot captures visual images. The robot's arm joints, which include ShoulderRoll, ShoulderPitch, ElbowRoll, ElbowYaw, and WristYaw on both arms, are likewise given ten units. The Rein_RNN has been taught in this situation to forecast the data's potential states. The outcome of the angle of the joint subunits is transmitted back into the input layer on a following time step during the assessment stage following instruction, as well as into the robot as a motorised command. . This allows us to interpret the sequences produced by output units of joint angles as the robot acting autonomously in response to commands. Two cameras were used to record footage, one webcam was used to record speech, while every person had a turn-mounted headphone for recording audio. The robot's movements were also continuously recorded. The transcriptions of the audio recordings were afterwards utilised as the foundation for a quantitative study of the encounter. The 48 audio and video recordings, as well as the text of the transcriptions, were subjected to a grounded theory analysis for the qualitative analysis. The data collection for the grounded theory analysis was primarily focused on the quantity and types of interaction in the three categories of "robot and individual learner," "robot and both learners," and "between the learners," as well as the quantity and kind of collaboration among the learners. The commonalities between the talks with the four distinct robot behaviours were then determined through data analysis in order to describe and compare them. The two students were handed tablets to complete an electronic questionnaire after each interaction with a single robot behaviour in which they were asked to rate the conversation on a Likert scale with a range of 0 to 5. To find out how learner preferences were influenced by robot behaviour, session order, learner gender, age, the second language level, experience with language restaurants, and routine, the evaluations regarding learning efficiency ('How could you assess the entire session? from an educator's perspective') and the robot's interaction behaviour ('How would you rate the robot's behaviour as a conversation participant?') were analysed.

- Storytelling model – Using stories to inspire kids to acquire a second language is an excellent strategy. In this model, certain short stories were altered in accordance with the course material. The teacher will make the following announcement at the start of this interaction: "Let robot tells a story for us." The robot then starts telling a story. The robot pauses after each brief paragraph to allow the teacher to quiz the kids on the story's themes. The robot would make an acclamation sound once the pupil provided an accurate response. After the asking action, the robot will proclaim the right response again.
- The Q and A model: In this model, a robot will randomly select a learner to approach it and ask him or her a few straightforward English questions. The robot will make an acclamation sound if the person answered the question correctly, and a ridiculous sound otherwise. The robot then requests that the student return to his or her seat.
- Cheerleader model: The instructor might arrange a competition game in class during which the students might be put into groups. The competition's events involve answering questions, choosing the appropriate word or image, and executing activities. The robot will dance or make a cheer sound for a learner when they respond correctly.
- Let's act model: In this framework, the English teacher will request in English that the pupils take certain actions, such as raising their hands, turning around, and moving somewhere. There are two pieces to this model. The teacher first gives the robot permission to carry out this command task. The robot gave the kids instructions and also carried out the actions. Second, the option to order the robot will be given to the students. A student will be chosen at random by the robot to do this activity.

The model of guiding pronunciation is the teacher letting a robot guide everyone as they speak English. As seen in figure 3, both the robot and the pupils concurrently alter their speaking rate and voice.



Figure 3. The robot is teaching to the students

Interaction with students and evaluating process

When teaching English, it is preferable to use technology to facilitate greater connection between teachers and students. Online instruction, online grading, and other approaches are effective applications of human-computer interaction. Multi-functional is gradually becoming an extremely essential supplementary tool in classroom instruction due to the enormous growth of information technology. The recent applications of multi-functional has a significant impact on the acquisition of learners: first, students use the Internet to gather and analyses various pieces of course-related material in preparation. The following describes the precise procedure for evaluating the effectiveness of instruction using principal component analysis. The initial English teaching performance assessment score combination is believed to be

$$Y = (Y_1, Y_2, Y_3, \dots, Y_p) \quad (18)$$

wherein P represents the total amount of indicators used to evaluate instruction. It is necessary to conduct standardized processing of collected data, and the standardized processing formula is

$$Z_{ab} = \frac{z_{ab} - z_b}{k_b} \quad (19)$$

One of them,

$$\begin{cases} z_b = \frac{1}{n} \sum_{a=1}^n z_{ab} \\ z_b = \frac{1}{n} \sum_{a=1}^n (z_{ab} - z_b)^2 \end{cases} \quad (20)$$

The linear regression matrix of evaluation metrics was calculated following the standardisation of instructional quality assessment criteria:

$$R = (corr_{ab})_{g \times g'} \quad (21)$$

where calculating technique $corr_{ab}$ denotes the coefficient of correlation between both the i-th instructional quality assessment sample and the j-th indicator.

$$corr_{ab} = \frac{1}{n-1} \sum_{k=1}^n z_{corr a} z_{corr b} \quad (22)$$

Create the characteristic curve $u=Ru$, and then determine its eigenvalue and eigenvalues., and then determine its eigenvalue and eigenvalues.

$$\gamma = (\gamma_1, \gamma_2, \dots, \gamma_p) \quad (23)$$

$$u = (u_1, u_2, \dots, u_p) \quad (24)$$

Determine the percentage of the accumulated variance that the key elements of the English teaching quality assessment index contribute to:

$$\beta = \sum_{i=1}^p \alpha_i \quad (25)$$

The assessment framework is quite straightforward and the evaluation procedure is fairly uncomplicated; the significance of the assessment model is calculated using the conventional statistical approach. Because the assessment method is very homogeneous and cannot be updated dynamically or have a lot of data added, the evaluation model is fixed and is not self-adaptive.

Performance analysis

Class question categorization results evaluation uses accuracy, mean evaluation time, interest of students (yes,no,undesirable),teaching modes (large impact, moderate impact and no impact).

Table-1 Confusion matrix for Rein_RNN evaluation

	Teacher’s question type- Predict the number of correct evaluations		
The actual number of correct evaluations	Grammar knowledge	General knowledge	Knowledge in novel
	1475	1875	1768
	164	157	138
	35	65	46

Table-2 performance on interest of students

interested	English reading (%)	English listening (%)	English writing (%)	English expression (%)
yes	98.6	89.6	79.6	71.5
No	97.6	89.6	78.6	72.5
Undesirable	94.4	89.5	83.2	77

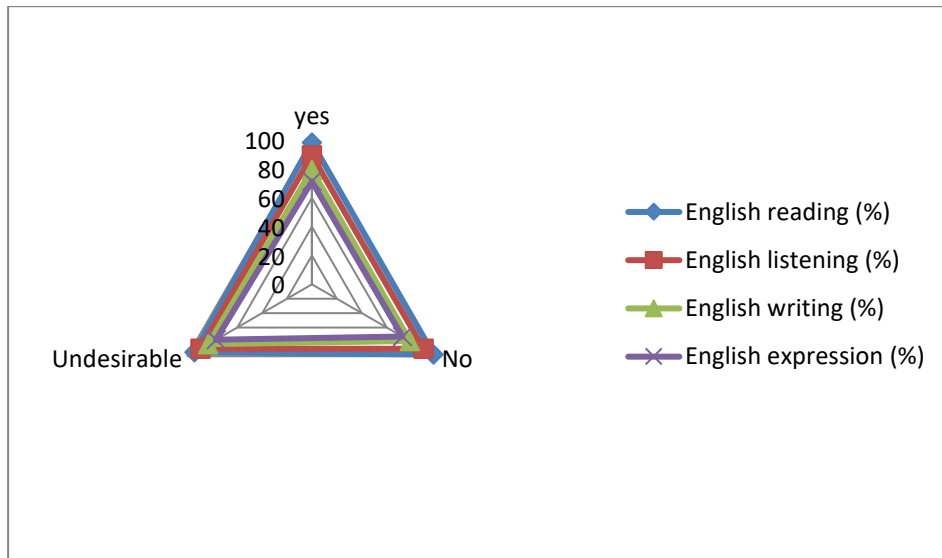


Figure-3 overall effectiveness about performance on interest of students

Figure 3 displays the overall effectiveness about performance on interest of students evaluation results in relation to the parameters of the teacher's questionnaire design. For English reading 98.6% of students are interested, English listening 89.6% of students are interested, 79.6% of students are interested in English writing and 71.5% of students are interested in English expression.

Table-3 Performance of teaching modes

Impact level	Hearing modalities (%)	Visual and tactile modalities (%)	Speech modalities (%)	Language modalities (%)
large impact	90.4	89.8	78.9	83.6
moderate impact	91.4	89.7	78.4	84.2
no impact	91.5	87.6	77	85.6

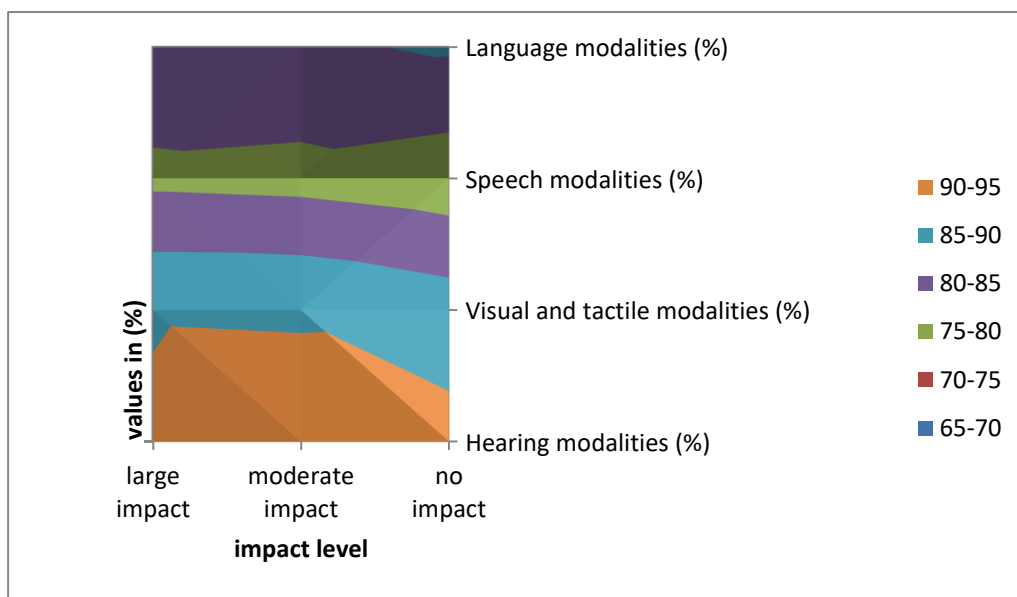


Figure-4 Performance of teaching modes favored by students

Figure 4 displays the parametric evaluation of teaching modes favored by students, where the impact level and percentage are displayed on the X and Y axes, respectively. While analyzing using Rein_RNN model on large impact, 90.4% on hearing modalities, 89.8% on visual and tactile modalities, 78.9% on speech modalities, 83.6%

on language modalities. In terms of moderate impact, it achieves 91.4%,89.7%,78.4% and 84.2% for the above mentioned modalities.

Table-4 Comparison of Rein_RNN model

Question type	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
novel type	97.6	89.6	76.4	81.4
Prompt type	95.7	87	78	83.4
grammar type	98	87.5	79.5	82
Application type	97.4	88	76	85.6
Evaluation type	95	85	75	86

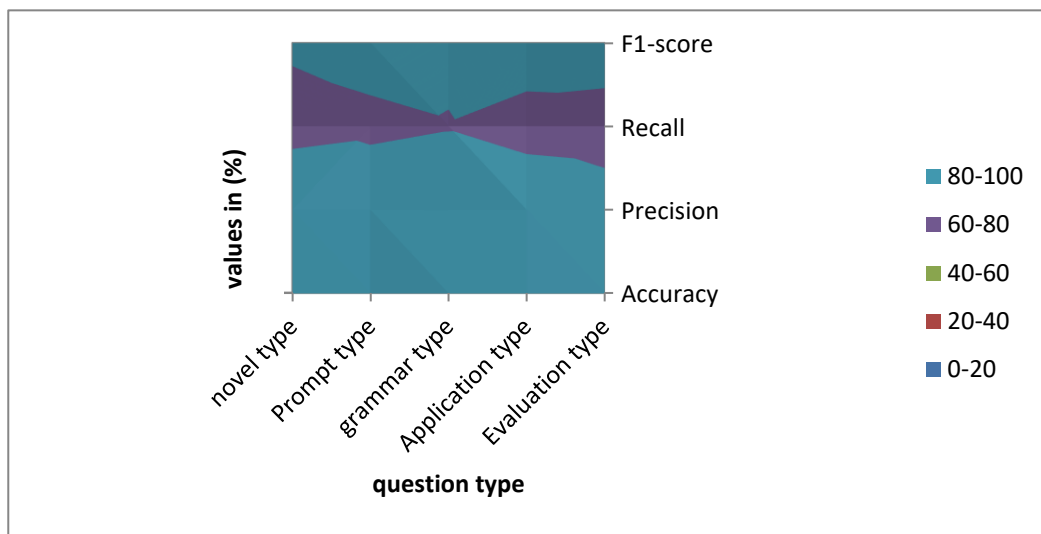


Figure-5 analysis of parameters for Rein_RNN model

Figure 5 displays the parametric analysis for the Rein_RNN model. Numbers for the question type and percentage are displayed on the X and Y - axis, accordingly. The Rein_RNN model, on the other hand, provides the highest accuracy, with 97.6% for novel type and 97.4% for application type. Precision for the novel type is 89.6%. The best recall value for grammar types is 79.5%, and for evaluation types, the F1 score is 86%.

Table-5 Comparison of various parameters with existing methods

Question type	Accuracy (%)	Mean evaluation time (sec)
Fuzzy neural network [10]	89	11
LSTM [13]	91	6.9
artificial neural networks [14]	79	5.9
Convolution Neural Network [15]	95	10
Rein_RNN	98.6	4.3

IV. CONCLUSION

Instead of conducting a true experiment with controls, this study was a field trial. To aid the teacher in teaching and encourage students to participate in learning activities, we introduce a humanoid robot into an English lesson. The majority of the pupils in the three classes had favourable attitudes towards this robot and were very interested in how it performed. Although first the teacher can be unsure of the robot's performance and unsure of how to

work with it. Our interaction models serve as a good illustration of this design. It is crucial to inform the teacher of the robot's capabilities and limitations. This article examines an artificial intelligence-based interactive multimodal teaching approach to English. Furthermore covered is the evaluation methodology for various modalities in relation to artificial neural networks. This demonstrates the need to integrate multimodal theory and recurrent neural networks into English teaching interactions from both theoretical and practical perspectives. This study shows that Reinforced Recurrent Neural Network-based Multimodal English Teaching Interaction Theory plays a part in improving the capacity of both educators and pupils to interact. This can assist students develop even further, foster a positive English-learning atmosphere, and enhance students' English-learning skills. The new educational equipment has not fully highlighted the human-computer interaction, and this demands more time and effort from teachers. In the future, educational institutions should invest more in HCI hardware while also focusing on teacher training programs to improve their HCI teaching skills and assist students acquire the language more efficiently.

REFERENCE

- [1] Alanis. Electricity prices forecasting using artificial neural networks. *IEEE Latin America Transactions* . 2018;**16**(1):105–111. doi: 10.1109/tla.2018.8291461. [[CrossRef](#)] [[Google Scholar](#)]
- [2] 2. Santosh T. V., Vinod G., Saraf R. K., Ghosh A. K. H. S. Application of artificial neural networks to nuclear power plant transient diagnosis. *Reliability Engineering & System Safety* . 2007;**92**(10):1468–1472. doi: 10.1016/j.res.2006.10.009. [[CrossRef](#)] [[Google Scholar](#)]
- [3] 3. Hodo E., Bellekens X., Hamilton A. Threat analysis of IoT networks using artificial neural network intrusion detection system. *Tetrahedron Letters* . 2017;**42**(39):6865–6867. [[Google Scholar](#)]
- [4] 4. Safa M., Samarasinghe S., Nejat M. Prediction of wheat production using artificial neural networks and investigating indirect factors affecting it: case study in Canterbury province, New Zealand. *Journal of Agricultural Science and Technology A* . 2018;**17**(4):791–803. [[Google Scholar](#)]
- [5] 5. aard T. K. Rhizomes and plateaus: a study of digital communities of practice in University College English Teaching. *Afore* . 2017;**55**(1):35–45. [[Google Scholar](#)]
- [6] 6. Hee Y., Jeong J.-Y. The effects of professor presence and interaction on PAD and satisfaction in a university class. *Journal of the Korea contents association* . 2017;**17**(7):144–157.
- [7] Jo, D.; Kim, G.J. ARIoT: Scalable augmented reality framework for interacting with Internet of Things appliances everywhere. *IEEE Trans. Consum. Electron.* 2016, 62, 334–340.
- [8] Karasinski, J.A.; Joyce, R.; Carroll, C.; Gale, J.; Hillenius, S. An Augmented Reality/Internet of Things Prototype for Just-in-time Astronaut Training. In *Virtual, Augmented and Mixed Reality*; Lackey, S., Chen, J., Eds.; Springer International Publishing: Cham, Switzerland, 2017; Volume 10280, pp. 248–260
- [9] Alam, M.F.; Katsikas, S.; Beltramello, O.; Hadjiefthymiades, S. Augmented and virtual reality based monitoring and safety system: A prototype IoT platform. *J. Netw. Comput. Appl.* 2017, 89, 109–119.
- [10] Pérez-Marín, D., Paredes-Velasco, M., & Pizarro, C. (2022). Multi-mode Digital Teaching and Learning of Human-Computer Interaction (HCI) using the VARK Model during COVID-19. *Educational Technology & Society*, 25(1), 78-91.
- [11] Premalatha, R., & Aiswarya, R. S. (2022). Physical education system and training framework based on human-computer interaction for augmentative and alternative communication. *International Journal of Speech Technology*, 1-11.
- [12] Zhang, L., Wei, P., Zhang, Y., & Wang, N. (2022). Artificial Intelligence and Edge Computing Technology Promote the Design and Optimization of Flipped Classroom Teaching Models for Higher Vocational, Ideological, and Political Courses. *Mobile Information Systems*, 2022.
- [13] Jing, Z. H. A. N. G. (2021). Application of Mind Map in Ideological and Political Teaching in Higher Vocational Colleges. *The Theory and Practice of Innovation and Entrepreneurship*, 4(9), 171.
- [14] Zhang, B., Velmayil, V., & Sivakumar, V. (2021). A deep learning model for innovative evaluation of ideological and political learning. *Progress in Artificial Intelligence*, 1-13.
- [15] Chen, X. (2021). Compensated fuzzy neural network-based music teaching ability assessment model. *Computational Intelligence and Neuroscience*, 2021.
- [16] Qin, R. (2021, June). Interactive Effect Evaluation Model based on Neural Network in the Guiding of Online ideological Education. In *2021 5th International Conference on Trends in Electronics and Informatics (ICOEI)* (pp. 836-839). IEEE.
- [17] Liu, Q. (2022). Research on the Innovation of Ideological and Political Education in Universities Based on Artificial Intelligence Machine Learning in a Wireless Network Environment. *Security and Communication Networks*, 2022.