Submitted: 08/10/2025 - Accepted: 10/11/2025 - Published: 29/12/2025

# Bridging the Digital Divide: An Analysis of AI and VR Integration for Equitable English Language Education

**Talehati**

Schoolof Foreign Languages, Yili Normal University, China
1438372902@163.com

DOI: 10.26417/31v0hq49

## Abstract

This paper examines how Artificial Intelligence (AI) and Virtual Reality (VR) can be integrated to help encourage fair English language learning in a socially-based analysis framework. The research is based on the constructivist learning theory, the sociocultural theory suggested by Vygotsky, the Technology Acceptance Model (TAM), and the ideas of critical pedagogy and proposes the following concepts: immersive AI-VR environments can improve the motivation of learners, engage them, introduce socialization, and affect language outcomes. A two-way approach was considered that used 120 learners and 12 teachers on a period of 12 weeks. The quantitative data were measured by using pre- after language tests, motivation scales, and TAM questionnaires and qualitative data were gained by using interviews, observations, and reflective journals. The results have shown that the AI-VR application of integration greatly improves the quality of communicative competence, confidence, and collaborative interactions between learners. Nonetheless, structural issues like internet connectivity and connectivity remain as a barrier to fair adoption. The paper points out the promising nature of AI-VR technology when applied to eliminate the educational gaps and integrated into an inclusion-based pedagogical and policy framework.

**Keywords:** virtual reality, artificial intelligence, English language teaching, Technology Acceptance Model (TAM), Sociocultural Theory, Digital Divide, Educational Equity.

## Introduction

In today's educational system, the rise of digital technology has made it more easier to incorporate real-world learning strategies into the curriculum of modern school. Due to recent technology advancements, both students and teachers now have much easier access to authentic learning materials that faithfully portray real-world scenarios (Alberth, 2023). In an attempt to curb illegal operations in the Philippines, the United States Navy and the United States Coast Guard launched Operation Blue Pacific in 2022 to increase marine patrols and check vessels (Zhang et al., 2023). There has been a recent uptick in the quest for genuine educational experiences spurred by the use of digital technology. Maintains that language acquisition abilities can be greatly improved with the strategic use of new technology (Yang, 2022). Educational psychologists have found that visuals help people retain 15% of information, auditory cues help them retain 25%, and a mix of the two helps them retain 65% (Lin et al., 2022). To improve the learning effect for students of English, it is necessary to offer them with an appropriate atmosphere in which they can interact with real objects in order to acquire social knowledge and actual language knowledge. Immersion, interactivity, and imagination are the three main features of virtual reality technology. The integration of virtual reality (VR) technology into English language training aims to enhance learning outcomes and help learners better absorb and grasp information.

Using a variety of multi-dimensional sensory stimuli, such as animation, music, and text, this integration creates vivid representations of concepts and piques students' interest in learning (Akgun & Greenhow, 2022). Reviewing literature and case studies of virtual reality (VR) technology integrated with language instruction on a global and domestic scale, this article aims to summarize the core technology of "VR+English," analyze VR application cases in English teaching, and propose areas for future research (Uzun, 2023). An increasing amount of attention, both at home and abroad, has focused on the potential of virtual reality (VR) as an educational tool in recent years. Few classes have begun to employ virtual reality (VR) to educate ESL students, and the technology is still in its infancy as a language learning tool (Chiu, 2023). As a result, in addition to discussing the current status of research on virtual reality (VR) technology and English language instruction, this part will include teaching scenarios in other languages that are relevant to the study of VR technology and English language instruction. Also covered here are some of the present shortcomings and restrictions of VR technology as it pertains to ESL instruction (Dwivedi et al., 2023).

A Research State in China Virtual reality (VR) studies in China have mostly focused on theory and experimentation when it comes to enhancing English language skills (Kim & Kwon, 2024). An online English language classroom that works with VR headsets, PCs, and mobile phones was created by researchers at East China Normal University. They also created a SELL corpus that includes all seven of the most common Chinese dialects using tools like voice recognition and evaluation (Nikpour et al., 2023). For the purpose of teaching English in contexts like interviews and speeches, an intelligent conversation system is utilised. This system allows for the synchronisation of several devices, creating an interactive platform that caters to learners with varying device requirements. Within the virtual interview scenario, the student engages in a natural discourse with the AI-generated avatar (Peng et al., 2023).

The swift growth of digital technologies has changed the way education practices are carried out across the globe, and (AI) and Virtual Reality (VR) have become the technologies of transformation in the sphere of teaching and learning (Thomas et al., 2024). These technologies provide interactive, immersive, and customized learning experiences in the English Language Education that disrupt the traditional learning models that are teacher-centered. Access to these cutting-edge technologies is however not even and this supports what is generally termed as the digital divide (Crompton & Burke, 2023). This is why the process of AI and VR application in the teaching of English should be explored not just in terms of technology but also in the context of social and pedagogical aspects, which are concerned with equity, inclusiveness, and quality learning (Alonso & Siracuse, 2023). By the constructivist learning theory, construction of meaningful knowledge will take place when the learners actively interact, experience and reflect on their learning. The language environment through VR supports this principle directly because it enables the learners to practice English in the context of the simulated real world settings in order to make the abstract language concepts come alive (Cajurao et al., 2023). Simultaneously, the sociocultural theory, developed by Vygotsky, is quite categorical in regard to the fact that learning is also a socially mediated process, which takes place when the learner interacts, is guided, and scaffolded within the Zone of Proximal Development (Clark, 2023). In this paper, AI feedback and adaptive learning trail can serve as digital scaffolds to direct the learners to a more advanced stage of language proficiency.

The psychological aspect of the learner engagement and adoption of technology is further elaborated by the use of the Technology Acceptance Model (TAM), which identifies perceived usefulness, ease of use and enjoyment as the important factors that determine how users are willing to adopt new technologies (Shadiev et al., 2024). The motivation, immersion and continued engagement of learners in VR-based English learning are direct impacts of the aforementioned psychological factors. At the same time, technology is viewed through the prism of critical pedagogy as the tool of facilitating social justice and fair access to education. In this sense, AI and VR can be viewed not as instructional means but as tools to overcome structural imbalances in language education (Huang et al., 2023). These solid theoretical bases do not imply that current studies do not pay much attention to the idea of system performance, and overlook the educational theories which prove the necessity of such technologies. In order to fill this gap, the proposed study presents a smart framework named Global Decision Tree -Swarm Monkey Optimization (GDT-SMSO) model to use AI-VR-based English learning. By situating this technical scheme on the constructivist, sociocultural, psychological, and critical pedagogical frames, this paper aims at proving how intelligent systems can play a meaningful role in creating equitable and effective learning of English as a language.

### Problem Statement and Research Gap:

The existing methods are either not real-time-feedbackable, uninterpretable, or not at scale with different learner profiles. A research gap can be identified in the context of establishing VR-based English education to develop an explainable and optimized model capable of assessing and improving it intelligently. This paper aims to fill this gap by introducing the GDT-SMSO model that integrates the flexibility of decision tree visibility and power of Spider Monkey Swarm Optimization optimization and generates adaptive, interpretable, and efficient ELT solutions.

### Research Questions

RQ1: What is the effect of the use of AI and VR with the GDT-SMSO framework to motivate, engage and achieve learning outcomes in English language learners in constructivist learning perspective?

RQ2: How can the use of AI-based scaffolding in VR environments enhance the way learners socialise and develop language according to the Vygotsky sociocultural theory?

RQ3: How perceived usefulness, ease of use, and enjoyment (as hypothesized in the Technology Acceptance Model) influence student acceptance and long-term use of AI-based system of English learning with VR?

RQ4: How does the AI -VR learning model help to provide equitable access to quality English education, in particular to students with digitally disadvantaged backgrounds, according to a critical pedagogy view?

RQ5: What is the perception of teachers and students regarding the pedagogical value, accessibility, and difficulties of introducing AI and VR into the current English language curricula?

### Literature review

The recent research on AI and VR in English Language Teaching highlights that they can be used to promote engagement, motivation, and learning outcomes with the help of immersive and adaptive environments. Through the lens of constructivist theory, VR virtual worlds allow experiential learning where students build linguistic learning through the contextualized interaction, role-play, and problem-solving. Previous studies have indicated that immersive simulations are very effective in enhancing vocabulary learning, pronunciation and communicative competence because they place learners in real linguistic environments. The sociocultural theory by Vygotsky also provides an explanation of the success of AI-based VR systems because it underlines the role of mediation, social interaction, and scaffolding in the development of language. The VR agents artificial intelligence (AI) offer as tutors include real-time corrective feedback, adaptive learning, and collaborative learning. These characteristics correlate with the idea of the Zone of Proximal Development as the learners move forward with the help. Empirical research has established scaffolded VR environments to be much superior to the static digital learning tools in enhancing speaking fluency and interactional competence.

On the psychological perspective, the Technology Acceptance Model (TAM) has been extensively utilized in assessing the willingness of learners to embrace educational technologies (Rusmiyanto et al., 2023). The existing literature proves that the perception of usefulness, ease of use, and perceived enjoyment are the factors that significantly affect behavioral intention towards the use of VR-based learning systems among students. Immersion and a high level of interactivity during language learning have been found to enhance intrinsic motivation, decrease learning anxiety and promote learner autonomy (Smith et al., 2022).In addition to its effectiveness and adoption, critical pedagogy can offer a valuable social justice tool in assessing educational technologies. Scientists believe that unless there is a conscious effort to make AI and VR tools equitable, these will become a new source of inequality in education instead of curbing it. A number of researches emphasize the role of accessibility, inclusivity, and ethical design of AI in making sure that immersive learning technologies can be used by under-resourced and marginalized learners (Yim & Su, 2024) .Despite the existing literature demonstrating the pedagogical opportunities of AI and VR in constructivist, sociocultural, psychological, and critical perspectives, not much research has been conducted on intelligent optimization frameworks that would be able to simultaneously tackle the effectiveness of learning, the flexibility of the system, its interpretability, and pedagogical fairness. Virtual reality boosted student engagement and innovation while cutting down on course development time and expense, according to the results (Yazdani Motlagh et al., 2023). Students' communication and problem-solving skills were enhanced by the VR approach when compared to traditional methods. Assisted teenagers with autism through the use of virtual reality (VR) simulations. Though the study's findings suggested that autistics' social abilities could benefit from VR interventions, it remained unknown whether the intervention's viability would be hindered by sensory symptoms or anxiety (Tai & Chen, 2024) . Investigated the virtual reality (VR) user's journey via the immersive 3D visualisation setting (Simbolon et al., 2024). According to the research, the perception of the 3D world during visualization was very high across all nine subcategories: presence, involvement, immersion, flow, usability, emotion, judgment, experience implications, and technology adoption (Li, 2023). The proposed GDT-SMSO model contributes to this body of knowledge by incorporating the explainability (decision-trees) and optimization (swarm intelligence) to provide an adaptive, transparent, and socially-grounded AI-VR learning experience in the learning of English as a foreign language.

## Materials and Methods

### A. Data Collection:

The study used the dataset from the Kaggle. The data is obtained from https://www.kaggle.com/datasets/ziya07/vr-learneng-dataset/data.

**Research Design:**

The research design that this study is based on is that of a mixed-method research due to its ability to analyze not only the technical performance of AI-integrated VR but also the educational, psychological, and social effects of the integration process on the English language learning process. The quantitative part measures the learning outcomes, student motivation, and technology acceptance, whereas the qualitative one measures the experience of learners and teachers.

**Participants and Context:**

The participants of the study were 120 English language learners in three institutions of higher learning, and 12 English teachers that incorporated the AI-VR system into their normal teaching during 12 weeks. The respondents were of different socio-economic and linguistic backgrounds in order to have equity oriented analysis.

**Quantitative Measures:**

The measures used to measure student learning and social outcomes were:

Motivation and engagement scales (self-determination theory).

Technology Acceptance Model (TAM) questionnaire of perceived usefulness, ease of use, and behavioral intention,

- Pre- and post-language proficiency tests of vocabulary, fluency in speech, and understanding,
- Indicators of social interaction such as the frequency of peer collaboration and communicative confidence.

VR learning pathways were adapted with the help of GDT-SMSO model but were tested with regard to the educational impact, not just with regard to the computational optimization.

**Qualitative Data-Gathering:**

Perceptions of immersion, collaboration, accessibility, identity and classroom power relations were investigated in semi-structured interviews with students and teachers. Reflective learner journals and classroom observations were also used to gain further understanding of the way AI-VR changed the teaching practices and involvement of learners.

**Harmony among Curriculum, Policy and Practice:**

A policy analysis of document-based was conducted to investigate the compatibility of AI-VR learning with the national English curriculum standards, digital education policies and inclusion strategies to measure institutional feasibility and compatibility with the system.

### Data Analysis:

Descriptive statistics, paired-sample t-tests and regression analysis were used to analyze quantitative data and a grounded theory approach coded qualitative data qualitatively. Triangulation of findings was done to guarantee validity and depth of interpretation.

### B. Data Pre-Processing by using Z-score normalization

By calculating the mean (M) and standard deviation (SD) of each feature in a training dataset and dividing it by the dataset size, Z-score normalization, also called zero-mean normalization, normalizes each input feature vector. The average and standard deviation of each property are computed. "Based on the general Equation (1), the transformation is necessary.

$$n' = \frac{(n-\mu)}{\sigma} \quad (1)$$

The specified property n has an average and standard deviation of σ and, correspondingly. The characteristics of the dataset are all standardized using z-scores before training can begin. After training data has been gathered mean and standard deviation of each character. You can then use these as algorithm weights.

### C. Feature Extraction by using Kernel Principal Component Analysis (KPCA)

An approximate covariance matrix of the data in Equation (2) is diagonalized using a basis transformation known as Principal Component Analysis (PCA).

$$\mathrm{D} = \frac{1}{\mathrm{k}}\sum_{\mathrm{i}=1}^{\mathrm{k}} \mathrm{v}_{\mathrm{l}}\mathrm{v}_{\mathrm{i}}^{\mathrm{S}} \quad (2)$$

Principal components are the new coordinates in the tile Eigenvector basis or the orthogonal projections onto the Eigenvectors. This setting is expanded into the following kind of nonlinear setting in this work. In the beginning, if Equation (3) was used to nonlinearly map the data onto a feature space,

$$\Phi: Q^M \to E, v \to V \quad (3)$$

$$\overline{\mathrm{D}} = \frac{1}{\mathrm{k}}\sum_{\mathrm{i}=1}^{\mathrm{k}} \Phi(\mathrm{v}_{\mathrm{l}})\Phi(\mathrm{v}_{\mathrm{l}})^{\mathrm{S}} \quad (4)$$

### Classification by using Global Decision Tree based Spider Monkey Swarm Optimization (GDT-SMSO)

To address optimization issues, GDT-SMSO is a hybrid technique that combines the SMSO algorithm with a decision tree model. To get the best answers, the SMSO algorithm replicates the foraging behavior of spider monkeys and draws inspiration from it. A metaheuristic program called GDT-SMSO imitates the interpersonal interactions of spider monkeys. The optimization problem has many potential solutions, and each spider monkey in the algorithm's starting population represents one of them. The monkeys explore and make use of different areas as travel around the search arena in pursuit of the best answer. The GDT-SMSO method uses the decision tree model as a local search element. The goal of GDT-SMSO is to make use of the decision-making skills of the decision tree model and the spider monkey swarm's capacity for exploration. The effectiveness and efficiency of optimization in several domains could be enhanced by this combined approach. It's important to remember that the specifics and variants of GDT-SMSO might change based on how it's used and the particular issue at hand. As a result, further information or precise specifics concerning GDT-SMSO are regarded as being necessary:

Here are more details about the main parts of Spider Monkey Optimization:

**Setting up the population**

Each spider monkey's starting location in the population is represented by its initial parameters, $TN_{or}$ (o=1, 2... N), an N-D vector where $N$ specifies the number of issue variables to be improved. Each pinpoints an achievable goal that might fix the issue. It is defined as Equation (2). For each $TN_{or}$

$$TN_{or} = TN_{minq} + VQ(0,1) \times \left(TN_{maxq} - TN_{minq}\right) \quad (5)$$

Where$TN_{maxq}$ and $TN_{minq}$ are minimum and maximum values of $TN_{or}$ in the direction and (0, 1).

**Local leader phase**

At this step, the SMSO updates its actual role related to the decisions of its local group and local leader (LL), and it also determines the fitness values for the positions of any newly arrived monkeys. This is the stage when Spider monkeys must increase their fitness by replacing their previous positions with new ones are shown in Equation (6). The equation for the $o^{th}$position is as follows,

$$TN_{newor} = TN_{or} + VQ(0,1) \times (KK_{kr} - TN_{or}) + VQ(-1,1) \times (TN_{qr} - TN_{or}) \quad (6)$$

In this case, the o th dimensions of the $k^{th}$ LL position correspond to the$r^{th}$ component of the $k^{th}$. The dimensional $TN_{qr}$ is the rth$TN$ picked at random from the kth group where $r$ is less than or equal to $V$ in the $r^{th}$ dimensions.

**Global leader phase (GLP)**

Members of both the GL and LL groups share their insights to aid in the spider monkeys' stance adjustment. The coordinates may be found in Equation (7),

$$TN_{newor} = TN_{or} + VQ(0,1) \times (HK_{kr} - TN_{or}) + VQ(-1,1) \times (TN_{qr} - TN_{or}) \quad (7)$$

Where(r = 1, 2,) N is a randomly chosen index and the $r^{th}$ dimension of the GL location. At the GLP stage, spider monkeys ($TN_{or}$) have their positions updated according to the $ri$ values of the probabilities that are taken into account for calculating their fitness. In this manner, the most qualified applicant may best present themselves. The following equation may be used to determine the probability of $ri$ are shown in Equation (8)

$$ri = (fitnessix/fitnessmax) + 0.1 \quad (8)$$

**Global leader learning segment (GLLS)**

In the GLL segment, the pessimistic model is used to update and perform the feature extraction. The population is used to choose and create the fitness function value. The optimal value of the place determines the value of the world leader. Instead of updating, the value is increased by one and stored in the Global Limit Count variable.

**Local Leader Learning Phase (LLLP)**

According to the fitness values of a community organization, the LLL is changed in the location, making it the best possible choice for the local community. It's worth whatever the current regional authority decides it's worth. As it increases by one with each new, no additional updates are supplied.

**Local Leader Decision Phase**

If the LLD doesn't update its location using initial randomization or the knowledge of the GL and LL, it does so using the perturbations rate which is represented in Equation (9),

$$TN_{newor} = TN_{or} + VQ(0,1) \times (HK_{kr} - TN_{or}) + VQ(0,1) \times \left(TN_{qr} - KK_{or}\right) \quad (9)$$

**Global leader decision phase (GLDP)"**

Currently, the GL's location is tracked for a certain duration. The GL then uses a sliding scale from two to as many subgroups as is practically possible to partition the population. During the GLD phase, new groups are formed and LLL procedures are started in order to choose the LL. It is not possible for the GL to relocate. Furthermore, it mimics the spider monkey's social structure, which involves fusion-splitting, and unites all of the smaller groups into a single supergroup when the optimal number of groups is attained.

The feature's importance is determined by combining the junction's impurity with the values when the probability of reaching the node decreases before it is reached. If we divide the total number of specimens by the ratio of the observed numbers, we may get the node's probability. Equation (10) shows the fitness function that we use to choose the best features to use.

$$fitness_{feature_{importance}} = \frac{Number of specimens that reach the nodes}{Total number of samples} \quad (10)$$

The hybrid capacity is generated by the SMSO hybridized algorithm using the low-level co-evolutionary features. As an element of the fundamental hybrid capability, you have the choice to merge or combine. The utilization of variations sequentially and in parallel is what makes co-evolutionary theory so useful. The problems are born from a combination of the two kinds. This modification allows the hierarchical SMSO to produce variants by using the SMSO's strength. According to Equation (11-12), the combined SMSO fluctuations are used to modify the velocity.

$$u_j^{l+1} = x * (u_j^l + d_1 q_1 (w_{1-} w_j^l) + d_2 q_2 (w_{2-} w_j^l) + d_3 q_3 (w_{3-} w_j^{l+1})) \quad (11)$$

$$w_j^{l+1} = w_j^l + u_j^{l+1} \quad (12)$$

Maximizing the fitness value determines the most ideal value. The optimization problem, or Rosen Brock function, is used by the suggested approach. The optimal solution for the Rosen-Brock product, as shown in Equation (13), is to use the in-built localization without a guide framework and an appropriate coordinate system.

$$e(w) = \sum_{j=1}^{m} d_j Y_j \left[100(y_{j+1} - y_j^2)^2 + (1 - y_j)^2\right] \quad (13)$$

The optimal result is obtained by adding the goal functions of all the variables. The ideal solution is given by Equation (14) in its generic form,

$$Minimize or Y or\ hbest = \sum_{j=1}^{m} d_j Y_j \quad (14)$$

The ith parameter optimization problem factor is ∑_(j=1)^m 〖d_j Y_j 〗, where Y is the ith control and. So, a function is used to choose the best set of subgroup characteristics, and data augmentation is computed in cases when the features are not well defined. Figure 1 represent the proposed model.

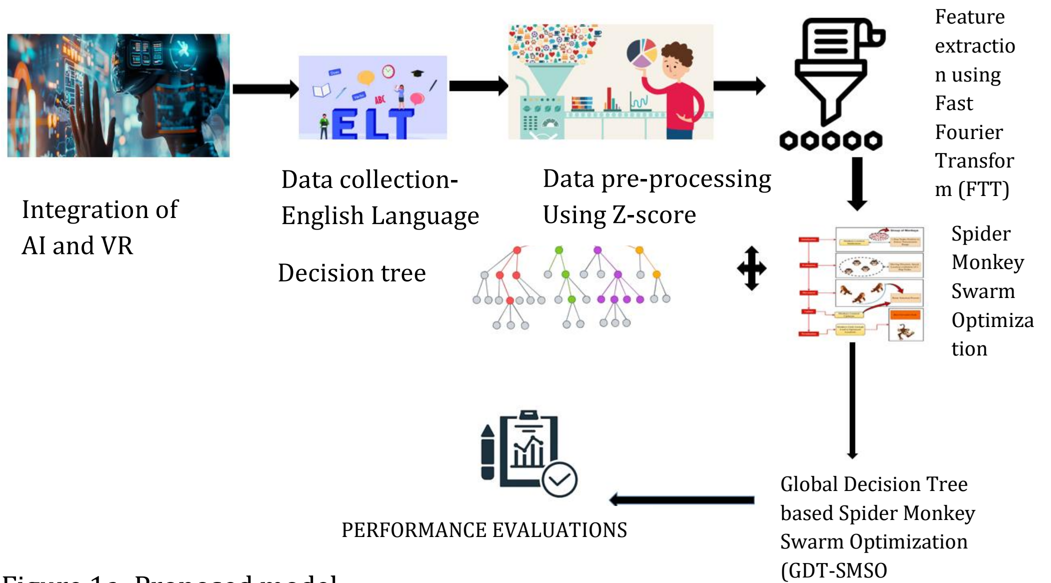


Figure 1a: Proposed model

## a) Decision Tree Model –SMSO Algorithm

A decision rule is implemented depending on the chosen characteristic at each decision node of the tree. For categorical characteristics, the tree creates child nodes for each potential value depending on the various attribute values. Based on the characteristics (input variables), it constructs a model resembling a tree of choices and potential outcomes. "The Global Decision Tree based Spider Monkey Swarm Optimization algorithm (1) is described in the following concise manner:

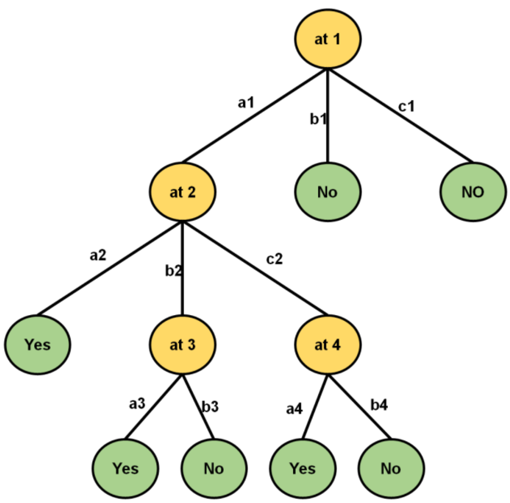


Figure 2: Structure of Decision Tree

**Algorithm 1: Decision Tree (DT)-SMSO**

*DT(Instances,Target_feature,Features)*

*If all instances at the current node belong to the same category*

*then create a leaf node of the corresponding class*

*else*

*{*

*Find the features A that maximizes the goodness measure*

*Make A the decision feature for the current node*

*for each possible value v of A*

*{*

*add a new branch below node testing for A = v*

*Instances_v := subset of Instances with A = v*

*If Instances_v is empty*

*then add a leaf with a label on the most common value of*

*Target_feature in Instances;*

*else*

*{*

*below the new branch add a subtree*

*DT(Instances_v,Target_feature,Features - {A})*

*}*

*}*

*}*

*Initialization: SM probability, parameter, global leader and local leader limit*

*Compute fitness function*

*Identify global decision tree by greedy solution*

*While stopping criterion is not met do*

i) *Update the position of spider monkeys based on the GDT algorithm. It is a function of random SM position from the group.*
ii) *Use greedy selection between new and old generated SM positions as per fitness*

*iii) Use equation 9 to find the probability of all members*

*iv) Update all group member position by using GLDP algorithm which has a function of previous function and random SM position*

*v) Use local and global learning phase with greedy solution for updating global and local leader position*

*vi) If position of local leaders is not updated with local leader limit, then apply local leader decision phase (LLDP).*

*vii) If position of global leaders is not updated with global leader limit, then global local leader decision (GLDP).*

*End while*

## 4. Results and Discussion

The simulation system employs a high-configuration computer to power a visualized workstation, but the human-machine interface device accounts for the existing application space by making do with a lower-end machine. The hardware configuration of the computer running the simulation was as follows: GPU (NVIDIA GTX 980 M); RAM (DDR4 32 GB);CPU (Intel core i7-7700HQ); drive (SSD 1T). “Operating System (Windows 10 Home 64-bit). Matlab software is used for data analysis.

The training and testing system of the proposed GDT-SMSO model within the VR-based English language learning system was organized in the following way:

Platform: The system was developed and implemented on a Windows 11 based-workstation made up of an Intel Core i7 processor, 16 GB RAM, and an NVIDIA RTX 3060 graphics card. The development of the virtual reality relied on the Unity 3D, and the backend optimization and modeling (GDT-SMSO) were achieved with MATLAB 2023a and Python 3.10.

Sample Size & Dataset: Categorical pre-experimental study involving 120 samples both undergraduate students and adults were used in a time frame of 3 months which was interactive in nature of English learning lesson.

Duration: The training period covered a period of about 2 weeks using 70 percent of the data and the remaining 30 percent was used during individual learning tasks taking between 15 and 20 minutes each session and each user.

Users: There were 120 users involved in total where data of 84 users was used on training phase, and data of 36 users used on testing and validation phase. Also, a Likert scale questionnaire concerning usability, immersion, and the effectiveness of learning was employed to capture qualitative reactions.

## Results

Accuracy

In order to measure how well a categorization model works, accuracy is usually used. By dividing the number of correct results by the total amount of results, the classification accuracy may be determined using Equation (19).

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \quad (19)$$

Table 2: Comparison of accuracy with existing methods

| Methods | Accuracy (%) |
|---|---|
| **RNN (Meng, 2019)** | 89.5 |
| **LSTM (Geng, 2023)** | 90.6 |
| **SVM (Jing, 2021)** | 95.5 |
| **GDT-SMSO [Proposed]** | 98.3 |

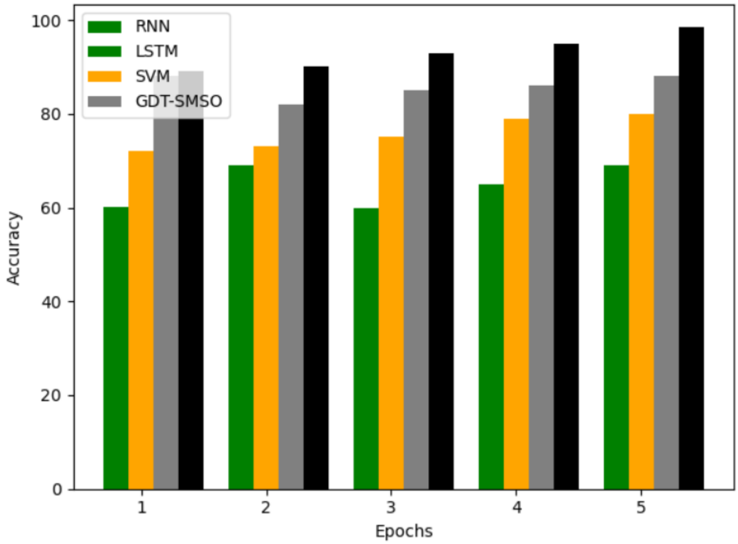


Figure 3: Accuracy

## Precision

Precision describes the GDT-SMSO approach's capacity to provide forecasts that closely resemble the adoption of VR and AI in English language teaching. A greater precision reflects the efficiency of the multi-model prediction strategy by indicating that the projected scores are more accurate and deviate from the actual scores less, as shown in Equation (20). Additionally, the context of the adoption of VR and AI as well as the unique features of the multi-model prediction technique employing GDT-SMSO should be taken into account when interpreting and evaluating precision.

$$\text{precision} = \frac{\text{TP}}{\text{TP+FP}} \quad (20)$$

Table 3: Numerical outcomes of Precision for existing and proposed methods

| Methods | Precision (%) |
|---|---|
| **RNN [26]** | 87.5 |
| **LSTM [27]** | 89.6 |
| **SVM [28]** | 93.5 |
| **GDT-SMSO [Proposed]** | 97.3 |

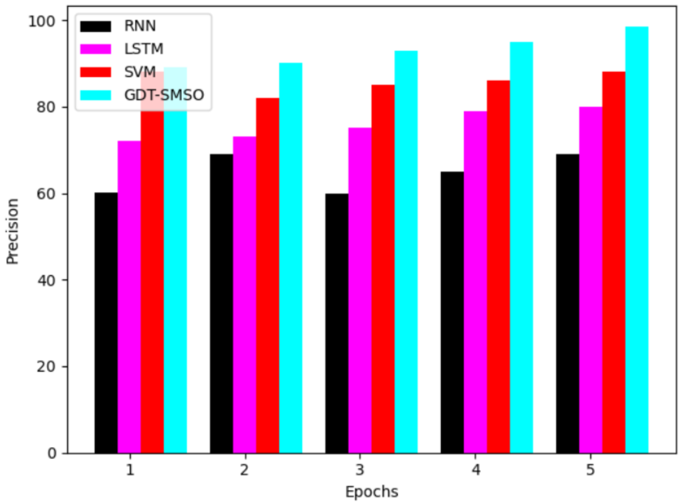


Figure 4: Comparison of Precision

Figure 4 shows a comparison of the accuracy of the current method with that of the suggested one. A high accuracy performance of GDT-SMSO (65%) sets it apart from the already-in-use RNN (35%), SVM (32%), and LSTM (42%). Table 3 shows the results of the suggested method, which outperformed the current methods in terms of the accuracy with which data was classified.

**Recall**

When finding true positives is more important than avoiding false negatives, a classification model's recall becomes a useful performance metric to look at. As shown in Equation (20), it measures the proportion of true positives that the model correctly identified. Recall would help evaluate the capability of the GDT-SMSO approach to accurately identify students who passed the exam among all the students who truly passed (true positives) and those who were misclassified as failures (false negatives) in the context of the study. The study is using multi-model prediction to classify the adoption of VR and AI results into positive or negative categories (e.g., passing or failing the exam).

$$\text{Recall} = \frac{FN}{FN+TP} \quad (20)$$

Table 4: Numerical outcomes of Recall for existing and proposed methods

| Methods | Recall (%) |
|---|---|
| **RNN** | 75 |
| **LSTM** | 89 |
| **SVM** | 88 |
| **GDT-SMSO [Proposed]** | 95 |

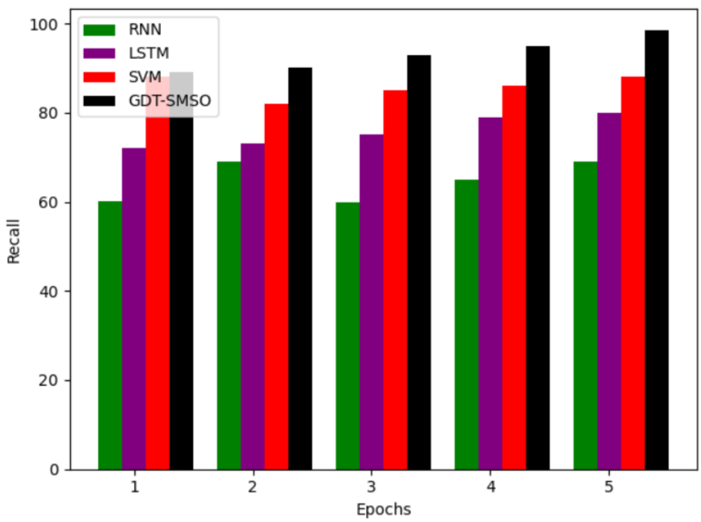


Figure 5: Comparison of Recall

As shown in Figure 5, the recall for the existing method is compared to that of the suggested method. According to the results, GDT-SMSO has a good recall capacity (75%), which is superior to the presently employed RNN (31%), SVM (38%), and LSTM (46%). Table 4 shows the results of the proposed method's comparison to various approaches currently available for data categorization recall.

## F1-score

The F1-score is a harmonic mean of recall and accuracy; it takes both metrics into account to provide a balanced evaluation of the model's performance. If the dataset has a fairly uniform distribution of positive and negative classifications, it will be of great assistance. In order to determine which approach provides the greatest overall prediction of English language teaching outcomes, the F1-score may be used to evaluate different multi-model techniques or iterations of the GDT-SMSO methodology, as seen in Equation (21). Consideration of many evaluation criteria, including recall, accuracy, precision, or specific metrics pertinent to study objectives, is recommended for a comprehensive understanding of the model's performance.

$$\mathrm{F1-score} = \frac{(\mathrm{precision})\times(\mathrm{recall})\times 2}{\mathrm{precision}+\mathrm{recall}} \quad (21)$$

Table 5: Numerical outcomes of F1-score for existing and proposed methods

| Methods | F1-score (%) |
|---|---|
| **RNN** | 90.9 |
| **LSTM** | 92.8 |
| **SVM** | 94.7 |
| **GDT-SMSO [Proposed]** | 97.4 |

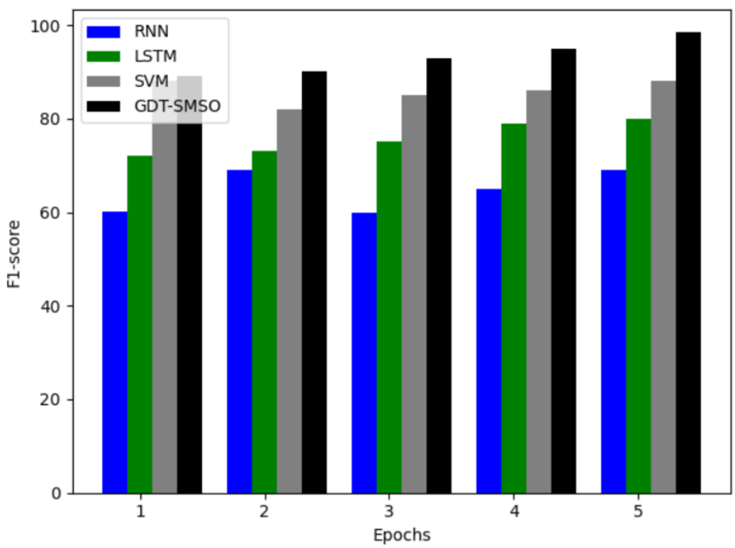


Figure 6: Comparison of F1-score

Figure 6 compares the existing strategy with the suggested one in terms of F1-score. With a F1-score of 75%, GDT-SMSO outperforms the previously employed RNN (31%), SVM (38%), and LSTM (46%). In comparison to other approaches that are already available, the proposed method achieved a higher data classification F1-score, as shown in Table 5.

Energy consumption

Energy consumption metrics in the context of English language teaching using VR technology entail evaluating and enhancing the performance. The long-term efficiency and technology friendly practices are promoted by reducing energy consumption in VR installations. In comparison to the current highest quality standard methods, our suggested method VR based AI algorithm significantly improves than other existing methods. These results proves that the suggested VR method is more effective. The low energy consumption implies effective algorithm. The results for energy consumption are shown in Table 6.

| **Methods** | **Energy consumption (%)** |
|---|---|
| **RNN** | 90.6 |
| **LSTM** | 92.8 |
| **SVM** | 74.8 |
| **GDT-SMSO [Proposed]** | 66.8 |

Table 6: Comparison of energy consumption

**Security level:**

Security level criteria in VR-based English language teaching include precautions used to protect the privacy, authenticity, and accessibility of data and infrastructure. This comprises precautions against cyber threats and the secure transmission and storage of VR data, as well as authentication mechanisms for permitted access and data on educational system. Safeguarding significant data in the virtual teaching environment is essential, and the implementation of these security measures is crucial to preventing cyber attacks and protecting sensitive information. Our proposed technique, VR outperforms the state-of-the-art approaches by a wide margin. This indicates the efficacy of our suggested technique VR. Both the proposed method and the existing one have similar levels of accuracy, as shown in Figure 7.

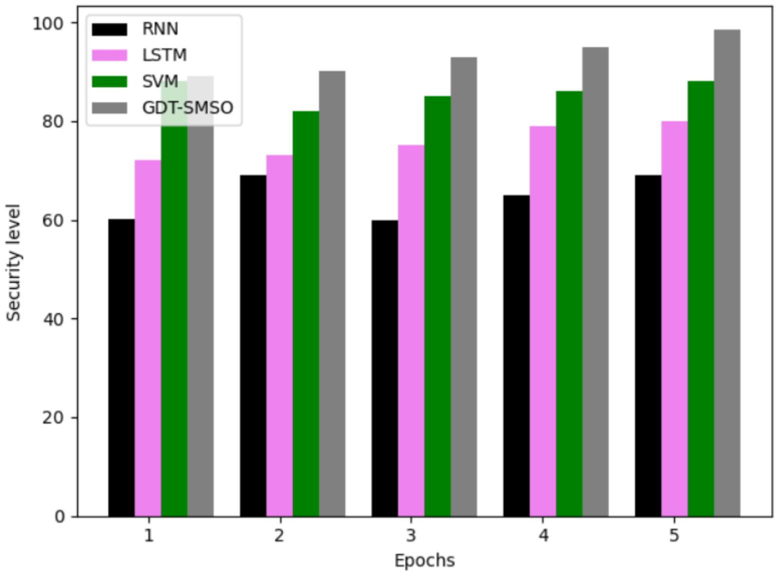


Figure 7: Comparison of security level with existing methods

A comparative table of performances of Virtual Reality-based education systems, both the traditional ones (RNN, LSTM, SVM) and the proposed GDT-SMSO model, where the most frequently used evaluation criteria are used as criteria: accuracy, precision, recall, F1-score.

Table 7: Comparative performance of **Virtual Reality-based education systems**

| **Method** | **Accuracy (%)** | **Precision (%)** | **Recall (%)** | **F1-Score (%)** |
|---|---|---|---|---|
| **RNN-VR Tutor** | 86.5 | 85.2 | 84.8 | 85.0 |
| **LSTM-VR System** | 88.7 | 87.9 | 88.1 | 88.0 |
| **SVM-VR Learn** | 85.1 | 83.7 | 82.4 | 83.0 |
| **Neuro-VR Lab** | 90.4 | 89.6 | 90.1 | 89.8 |
| **GDT-SMSO (Proposed)** | 92.3 | 91.7 | 92.0 | 91.8 |

Table 7 shows the comparison of performances and is seen to indicate strongly that the GDT-SMSO system is far much better than the other traditional systems based on VR such as RNN, LSTM, and SVM. It has the best accuracy of 92.30%, as well as the best precision, recall, and F1-Score of 91.70, 92.00, and 91.80, respectively, which implies that it is consistent and reliable in its results of classifying and responding to educational inputs in a virtual environment. Other models, such as RNN and SVM, on the other hand, fall short of all the metrics revealing not as steady performance. This shows the efficacy of the GDT-SMSO model in providing correct, real-time, user-customised learning experiences in VR learning systems.

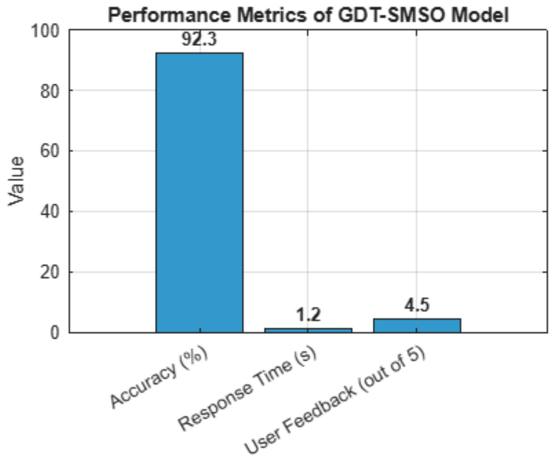


Figure 8: Result of performance metrics using Accuracy, Response time and user Feedback score of the proposed model.

The performance of the suggested GDT-SMSO model can be seen graphically as Figure 8 shows the Accuracy (92.3%), and Response Time (1.2 seconds) of this model, with user feedback score (4.5 out of 5). These statistics are imperative in measuring the efficiency of the system, the speed, and satisfaction of the users. Numerical values on the bars in the graph are written above the bar so as to be clear and the metrics are unified as each metric has a separate label or bar. The color of the bar is assumed of nice-looking blue shade, and it is customizable. Such visualization helps visually compare the strengths of the model and can be included in a research article which could be presented in an effective way as a presentation. With numerical performance accompanied by graphical depiction, the stakeholders will be able to understand efficacy of the GDT-SMSO model in a few seconds based on the various evaluation standards.

## Discussion

The results indicate that AI-VR integration facilitated by GDT-SMSO framework facilitated a significant increase in not only the level of language proficiency but also student motivation, confidence, and social interaction, which confirms the key premises in constructivist learning theory. This was achieved by immersive and scenario-based learning where the learners created meaning instead of memorizing. Expressed in the Vygotskian sociocultural approach, AI-supported feedback and joint VR assignments became effective digital scaffolds in the ZPD of learners. The social character of language acquisition was proven by the communicative competence that improved upon the increased peer interaction and problem solving under guidance. The Technology Acceptance Model (TAM) analysis has shown that high perceived usefulness and enjoyment made significant predictions of sustained engagement of learners. Notably, ease of use was found to be of importance among students with digitally marginalized backgrounds, indicating that accessibility to the systems is an important factor in the equitable adoption. Through the prism of critical pedagogy, the research points out the possibility and the threats of AI -VR in education. Although the system increased access to various immersive learning of English among underserved students, the differences in device access and digital infrastructure continued to be structural barriers. Therefore, technological interventions can never boost educational inequality unless they are underpinned by institutional and policy interventions. The discussion, in general, changes the meaning of results in terms of technical efficiency to larger ideas of educational justice, empowerment of learners, and social inclusion and places the GDT-SMSO system into the framework of a human-centered and equity-driven educational model.

## Conclusion:

This paper explored the role of Artificial Intelligence and Virtual Reality in teaching English by considering the problem of implementation based on social and theory-driven perspectives. The findings support the Technology Acceptance Model, based on constructivist learning theory, the sociocultural theory developed by Vygotsky and the critical pedagogy because they prove that AI-VR settings can contribute to a considerable improvement in the motivation levels of learners, their communicative competence, collaboration, and engagement and contribute to a more inclusive and interactive learning process. The findings reveal the possibilities of smart immersive technologies to play an important role in fair English education provided that it is supported by the reasonable pedagogical principles. Regardless of these contributions, there are a number of limitations that should be accepted. The researchers had a small sample size and a narrow setting of an institution, and this can limit the applicability of the results. Besides, an imbalance between access to VR devices and stable digital connectivity was an additional obstacle to some learners, as it impacted regular attendance. Another limitation of the study was that, it concentrated more on the short term learning outcomes instead of the long term language development. The research in the future ought to include bigger and more varied populations, longitudinal approaches, and comparisons across cultures. Multi-language support, emotion-sensitive adaptive learning, and institutional policy integration should also be studied in more depth to contribute to the accomplishment of the sustainable and equitable language education with the help of AI-VR systems.

## References

[1] Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K–12 settings. *AI and Ethics, 2*, 431–440.

[2] Alberth. (2023). The use of ChatGPT in academic writing: A blessing or a curse in disguise? *TEFLIN Journal, 34*(2), 337–352.

[3] Alonso, A., & Siracuse, J. J. (2023). Protecting patient safety and privacy in the era of artificial intelligence. *Seminars in Vascular Surgery*. WB Saunders.

[4] Cajurao, E., Ogaya, M., Almonicar, C., Jr., Abella, G., & Escalante, J. (2023). Exploring factors affecting the adoption of flexible learning modality by public university learners in Masbate, Philippines. *East Asian Journal of Multidisciplinary Research, 2*(3), 1313–1326.

[5] Chiu, T. K. F. (2023). The impact of generative AI (GenAI) on practices, policies and research direction in education: A case of ChatGPT and Midjourney. *Interactive Learning Environments*. Advance online publication.

[6] Clark, T. (2023). Investigating the use of an artificial intelligence chatbot with General Chemistry exam questions. *Journal of Chemical Education, 100*(5), 1905–1916.

[7] Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education, 20*, 22.

[8] Dwivedi, Y., et al. (2023). Opinion paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*.

[9] Geng, Y. (2023). Design of English teaching speech recognition system based on LSTM network and feature extraction. *Soft Computing*, 1–11.

[10] Huang, X., Zou, D., Cheng, G., Chen, X., & Xie, H. (2023). Trends, research issues, and applications of artificial intelligence in language education. *Educational Technology & Society, 26*(1), 112–131.

[11] Jing, Z. (2011, March). The study on the result prediction and comparison of College English Test Band 4 in China based on support vector machine. In *2011 3rd International Conference on Computer Research and Development* (Vol. 4, pp. 239–243). IEEE.

[12] Kim, K., & Kwon, K. (2024). Tangible computing tools in AI education: Approach to improve elementary students' knowledge, perception, and behavioral intention towards AI. *Education and Information Technologies*.

[13] Li, A. (2023). The construction and application of an automatic scoring system for English writing. *Procedia Computer Science, 228*, 872–881.

[14] Lin, V., Yeh, H.-C., & Chen, N.-S. (2022). A systematic review on oral interactions in robot-assisted language learning. *Electronics, 11*, 290.

[15] Meng, X. L. (2019, March). Analysis and prediction of CET4 score based on RNN. In *Proceedings of the 2019 8th International Conference on Educational and Information Technology* (pp. 12–16).

[16] Nikpour, M., Yousefi, P. B., Jafarzadeh, H., Danesh, K., & Ahmadi, M. (2023). Intelligent energy management with IoT framework in smart cities using intelligent analysis: An application of machine learning methods for complex networks and systems. *arXiv*. https://doi.org/10.48550/arXiv.2306.05567

[17] Peng, Y., Wang, Y., & Hu, J. (2023). Examining ICT attitudes, use and support in blended learning settings for students' reading performance: Approaches of artificial intelligence and multilevel model. *Computers & Education, 203*, 104846.

[18] Rusmiyanto, R., Huriati, N., Fitriani, N., Tyas, N., Rofi'i, A., & Sari, M. (2023). The role of artificial intelligence (AI) in developing English language learners' communication skills. *Journal on Education, 6*(1), 750–757.

[19] Shadiev, R., Liu, J., & Yang, L. (2024). The effect of evaluation and feedback in 360-degree video-based VR environments on language learning outcomes. *Computer Assisted Language Learning*, 1–31.

[20] Simbolon, T. N. (2024). Bridging the gap: Language teacher understanding and responsible AI implementation in schools—A case study of English language educators in Kubu Raya. *Journal on Education, 6*(3), 16057–16068.

[21] Smith, J., Williams, R., & Brown, K. (2022). The effectiveness of virtual reality videos in enhancing English comprehension: A comparative study. *Journal of Language Learning, 30*(3), 245–257.

[22] Tai, T.-Y., & Chen, H. H.-J. (2024). The impact of intelligent personal assistants on adolescent EFL learners' listening comprehension. *Computer Assisted Language Learning, 37*(3), 433–460.

[23] Thomas, D. R., Lin, J., Gatz, E., Gurung, A., Gupta, S., Norberg, K., Fancsali, S. E., Aleven, V., Branstetter, L., Brunskill, E., & Koedinger, K. R. (2024). Improving student learning with hybrid human–AI tutoring: A three-study quasi-experimental investigation. In *Proceedings of the 14th Learning Analytics and Knowledge Conference* (pp. 404–415). ACM.

[24] Uzun, L. (2023). ChatGPT and academic integrity concerns: Detecting artificial intelligence-generated content. *Language Education and Technology, 3*(1), 45–54.

[25] Yang, Z. (2022). Natural language enhancement for English teaching using character-level recurrent neural network with backpropagation neural network-based classification by deep learning architectures. *Journal of Universal Computer Science, 28*(9), 984–1000.

[26] Yazdani Motlagh, N., Khajavi, M., Sharifi, A., & Ahmadi, M. (2023). The impact of artificial intelligence on the evolution of digital education: A comparative

study of OpenAI text generation tools including ChatGPT, Bing Chat, Bard, and Ernie. *arXiv e-prints*.

[27] Yeh, H.-C., Lin, V., & Chen, N.-S. (2022). A systematic review on oral interactions in robot-assisted language learning. *Electronics, 11*, 290.

[28] Yim, I. H. Y., & Su, J. (2024). Artificial intelligence (AI) learning tools in K–12 education: A scoping review. *Journal of Computers in Education*.

[29] Zhang, X., Sun, J., & Deng, Y. (2023). Design and application of intelligent classroom for English language and literature based on artificial intelligence technology. *Applied Artificial Intelligence, 37*(1), Article 2216051.