Study on Artificial Intelligence (AI) Assisted English Linguistics Teaching System

Xin Meng(✉)

School of Ph.D. in English Language, Lyceum of the Philippines University, Manila, Philippines
xin.meng@lpunetwork.edu.ph

Abstract. Artificial Intelligence offers new solutions for English learning, and its emergence has broken the traditional way of learning English. This paper aims to investigate the Artificial Intelligence (AI) assisted English linguistics teaching system, conducting the numerical simulation to validate the performance. Based on intelligent English education and English linguistic theory, we have carried out three sets of models to test the influence of AI on English learning: (1) Face recognition algorithm is applied for the student concentration estimation. (2) Speech recognition is used to test students' pronunciation accuracy. (3) Cloud storage is applied to realize the efficient storage of the above algorithms. The findings from this research showed that we can design teaching content and methods in a customized chemical way by considering the students' unique English learning needs, and conduct comprehensive analysis through intelligent analysis of the students' background levels based on the artificial intelligence system. This research also contributes to expanding the development and application of Artificial Intelligence, as well as the improvement of front-line English teaching practice designs.

Keywords: Artificial Intelligence · English Linguistics · Teaching System · Face Recognition · Speech Recognition · Cloud Storage System

1 Introduction

The algorithm, which can directly promote the formation of artificial intelligence, is a key component of artificial intelligence. According to the historical development and research process of artificial intelligence, it can be seen that artificial intelligence [1–4] has gone through different stages of the development from its inception to the present. The application of artificial intelligence in education has received extensive attention. Artificial intelligence can use the collected data to make the corresponding activities and feedback, provide adaptive and personalized learning and personalized tutoring especially for the smart English guiding [5–7].

The AI for English guiding can be considered from the 3 core aspects.

Based on the artificial intelligence system, we can design teaching content and methods in a customized chemical way by considering the students' unique English learning needs, and conduct comprehensive analysis through intelligent analysis of the students' background levels, so as to enhance students' learning experience.

C. F. Peng et al. (Eds.): EIMT 2023, AHSSEH 8, pp. 511–520, 2023.
https://doi.org/10.2991/978-94-6463-192-0_67

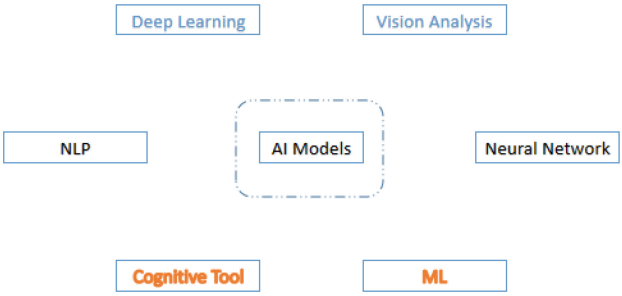

Fig. 1. The components of the AI

Relying on big data and massive information, artificial intelligence can track and record various data of learners. Then model, analyze and evaluate study results of learners, and effectively intervene on the learners.

The artificial intelligence word memory software [8, 9] can evaluate according to the individual situation of the students who use it, and formulate a learning plan suitable for their steadily increasing vocabulary.

The application of artificial intelligence has greatly improved the personalization of the English education in higher vocational colleges. According to the characteristics of the students, targeted teaching programs are formulated, effective learning methods are selected, and real-time analysis of students' learning situation is carried out by intelligent means, so as to flexibly adjust English teaching plans in higher vocational colleges. So, this paper studies the Artificial Intelligence (AI) assisted English linguistics teaching system and conduct the numerical simulation to validate the performance. To begin with, in the Fig. 1, we show the general components of the AI technology.

The designed model will be discussed in the following sections, along with the theoretical analysis. The rest of the sections are organized as the follows. In the Sect. 2, the literature review is discussed, we reviewed the latest studies on the intelligent English teaching and the basic models for the AI. In the Sect. 3, the proposed methodology is discussed considering the components of the face recognition, speech recognition and the cloud storage. In the Sect. 4, the numerical simulation on the performance is presented and finally, in the Sect. 5, we summarize the whole work.

2 Literature Review

The state-of-the-art intelligent English education and the English linguistics theory advancement are both reviewed to serve as the basis of the designed model.

For modern methods education of intelligent English language [10–12]. The application of the artificial intelligence in English teaching at higher vocational colleges has promptly innovated the English teaching model, meeting the demand for high-quality English teaching, quickly identifying and mading precise breakthroughs in the face of the difficulties and problems in modern English teaching. Artificial intelligence has obvious advantages in the teaching of spoken English. The emergence of educational robots has created a more colorful English communication environment for students,

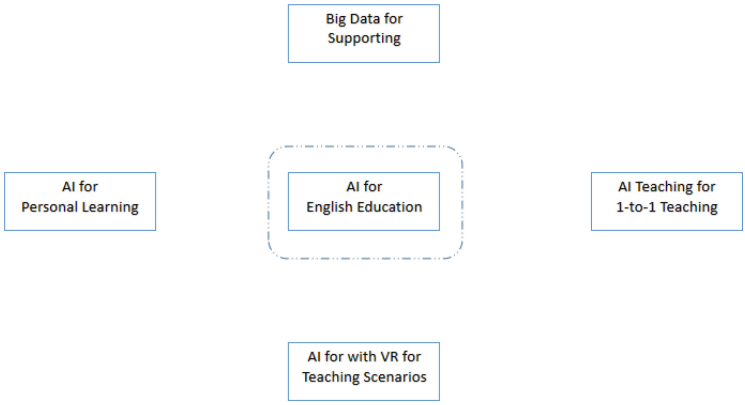

Fig. 2. The AI for the Intelligent English Education Methods

and accompanying dialogue is of great help for learners to become more accustomed to using English. With automatic feedback, learners can correct their own mistakes, which is extremely helpful for self-learners to learn and practice independently, so that they dare to open their mouths when learning English, and do not have to rely entirely on teachers, and become more independent in terms of learning content, learning style and learning time. Figure 2 shows the AI for the intelligent English education methods.

For the English linguistics theory [13–15] advancement. English is considered as a language subject, in order to produce achieve optimal teaching results, it must also cooperate with an appropriate teaching situation. By strengthening linguistics teaching, it helps students accurately master relevant English language knowledge, develop good habits of self-study and research, in order to properly apply what they have learned in real life.

3 The Proposed Methodology

The proposed model is discussed in this section which contains the following 3 components.

(1) Face recognition for student concentration estimation.
(2) Speech recognition used to test students' pronunciation accuracy.
(3) Cloud storage to realize the efficient storage of the above algorithms.

3.1 The Face Recognition Algorithm for Student Concentration Estimation

According to the expression, we can guess the other party's mental activities and emotions. Facial expression recognition is mainly divided into the three main steps: image acquisition and preprocessing, feature extraction and expression classification, among which feature extraction is a key part that determines the effect of facial expression recognition.

Facial information identification [16–18] should be considered in the beginning of experiments. The Local Binary Pattern is an operator used to describe the local features

Fig. 3. The CNN Structure Model

of an image, and it has significant advantages such as gray scale invariance and rotation invariance. For the local facial data set, the Facenet model is used to extract facial features, and the processed image information will obtain a multi-dimensional feature vector after machine learning.

The feature vector will be used as the basis for facial information matching and the CNN will be used for the face recognition. The input image is first convoluted through the convolutional layer and generates a specific type of activation feature map; then the space size of the feature map and the amount of network calculation are reduced through the pooling layer, and the nonlinear feature extraction ability of network is improved. In the Fig. 3, the CNN structure model is defined.

We use the VGG network model, it is better if we use multiple small convolution kernels. Non-linear operations can be realized through the activation function, and a better network structure can be trained without increasing the cost. The activation function chooses the RReLU function, which can make up for the shortcomings of the ReLU function defined as Eq. 1.

$$R\,\mathrm{Re}LU(x) = \begin{cases} ax, & x < 0 \\ x, & others \end{cases} \tag{1}$$

Where the a defines the randomly drawn values from a given uniform distribution, fixed during testing. In this paper, the network model adds an attention mechanism to the convolution of the last layer of VGG16, and reduces the number of the nodes in the subsequent two fully connected layers. The main purpose of this stage is to match and store the multi-dimensional feature vector obtained in the facial feature information extraction stage with its label in the local facial data set, so as to improve the speed and accuracy of subsequent facial feature matching. In the formula 2, we define the activation function.

$$x_L = x_l + \sum_{i=l}^{L-1} F(x_i, W_i) \tag{2}$$

Since each channel feature map is then convolved by its corresponding convolution kernel, the features learned by different channels contribute differently to the final expression classification. The network uses the RReLU activation function to deal with the nonlinear relationship between the input and output of the facial expression network. RReLU has its own sample normalization feature, which is not disturbed by external factors, and can solve the problem that the output of non-positive intervals is 0. When the

input is given a negative number, the operation of the neural unit can still be maintained. The feature vector after the fully connected layer is normalized by the Sigmoid function, the specific formula is defined as lemma 3.

$$\sigma(x) = \frac{1}{1 + \exp(-x)} \tag{3}$$

Where the x defines the input weight vector. Then, we use the transfer learning algorithm to achieve the recognition algorithm. Transfer learning needs to keep the convolutional layer structure of the feature extraction model unchanged, load the pre-trained weights and parameters, design a fully connected layer suitable for new tasks, and form a new network model with the original convolutional layer, and then use the new data set training model. We use the loss function to continuously optimize the network weight ratio to evaluate the quality of the network model, and then, the student concentration estimation will be achieved.

3.2 The Speech Recognition Used to Test Students' Pronunciation Accuracy

The artificial intelligence speech recognition engine is the core of the entire speech recognition system, responsible for recognizing speech and converting it into text sequences. Speech is input as data for the recognition process, which is converted from audio form to spectrogram form and passed to the engine. For the recognition, it contains the following cores. (1) Speech recognition front-end interface, we can input recorded audio files for speech recognition. (2) Collecting and managing voice materials in the different service areas, including data import and export, data annotation, and also data management, enables unified management and open sharing of voice materials and data. (3) Evaluate the model recognition speed, recognition accuracy, etc. according to the actual usage of the user.

The model we designed is divided into two parts: a generative model and a discriminative model. In the generative model, the text sequence is encoded and combined with randomly generated noise, and the speech is generated by the deconvolutional neural network, and then the generated speech is passed into the discriminant model. The intermediate value of a certain dimension is generated through the convolutional neural network, which is combined with the same text sequence encoding in the generation model. In the Fig. 4, we denote the speech signal peaks.

We need to find solutions to two major challenges for the designed model.

The impact of wrong words on the translation of the other words. When there is a speech recognition error in a sentence, analyze which words are likely to cause translation errors in other words. In order to analyze the impact of each wrong word on translation, make sure that there is only one word is wrong in each sentence during the experiment.

By comparing the original sentence with the speech recognition results, the error types are divided into homophone characters, personal names, homophone characters and more words with fewer words, and the proportion of them is calculated.

For the designed mode of the speech recognition, we consider the following steps. Firstly, the module is initialized; secondly, the module is compiled to identify the list; finally, the component starts to identify and finally generates a response interrupt. (1)

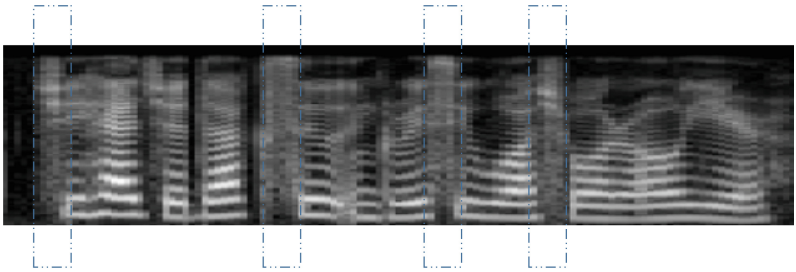

Fig. 4. The Speech Signal Peaks

Start module initialization. Perform operations such as reset, mode setting, etc. (2) The module compiles the identification list. Each identification entry corresponds to a 1-byte specific numerical value, which can be repeated, but the value cannot exceed 256. It is written into the module in the form of pinyin strings, and every two pinyin must be separated by spaces. (3) Parts begin to identify. Digital-to-analog gain represents the volume of the microphone. The higher the value, the higher the microphone volume, the more sensitive the speech recognition, and the higher the possibility of error.

Then, the students' pronunciation accuracy can be achieved.

3.3 The Cloud Storage to Realize the Efficient Storage of the Above Algorithms

Cloud storage [19–22] emerges as the new network storage technology. While enjoying the convenient services of cloud storage, users are also concerned about data security issues.

In the designed model, we consider the MongDB. The MongDB database is an open-source database system based on distributed file storage. In the case of high load, adding more nodes can ensue the performance of the server. It stores data as a document, and the storage data structure consists of key-value pairs. The database system used in this system is MongoDB, which is a document database with the advantages of low storage cost and high flexibility.

The document-based data structure is relatively more suitable for this system, especially in the storage and docking of unstructured data such as file meta-information. For the storage system, we consider the 3 core aspects.

Each data owner has a private cloud server. The "honest and trustworthy" private cloud server receives the plaintext data and the network model and trains it individually. After each round of training, the parameter results are federated with the parameter server and the trained neural network model is returned to the data owner.

After receiving the request trap sent by the parameter server, the relevant encrypted file is found out through calculation and sent to the data user.

While ensuring the confidentiality and availability of outsourced data, we should also study its unavailability, that is, how to safely delete data, make it invalid and unrecoverable, and ensure that data is "unavailable" when it is "not used".

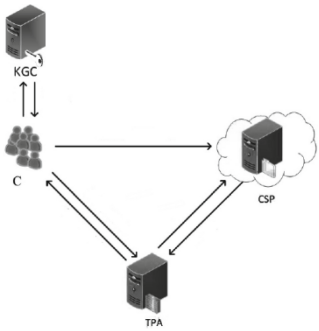

Fig. 5. The Cloud Storage System Structure

In the essential aspect, the cloud storage should be referred, the coding coefficient vector corresponding to the coding block generated after repairing according to the DPR restoration scheme is linearly independent from the coding coefficient vector corresponding to the undamaged coding block. For the vector group A, the necessary and sufficient condition for each vector in the vector group to be linearly independent is if and only if the parameter is 0 as lemma 4.

$$\sum_i^n \phi_i \mu_i = 0 \tag{4}$$

When repairing the cloud storage model, all row vectors in the repair matrix participate in the repair process, that is, the (n-m) row vectors and (n-1) repair data blocks in the (n-m) × (n-1) order repair matrix perform matrix operations to obtain a new (n-m) coding blocks, then n-m new coding blocks and coding blocks on undamaged storage nodes are linearly independent. In a centralized environment, a trusted server is used as a third party to manage and delete keys. This type of the scheme is not suitable for large-scale and dynamic user scenarios, and it needs to be assumed that the third-party server is trusted, hence, the model will be achieved, and the Fig. 5 shows the structure of the storage system. With this system, the algorithm will be stored and the (AI) assisted English linguistics teaching system will be achieved.

4 Experiment and Simulation

The experiment will be focused on the cloud storage system for the (AI) assisted English linguistics teaching system.

For the storage model, we consider listed steps.

Data integrity. When a user applies for using data stored on the cloud server, the cloud service provider must prove to the user that the stored data is completely preserved and available.

No block verification. Both user and third-party information can verify the integrity of the data stored in the cloud without storing file data.

Block insertion. The server can insert anything on existing client's files, or bring in new client's files.

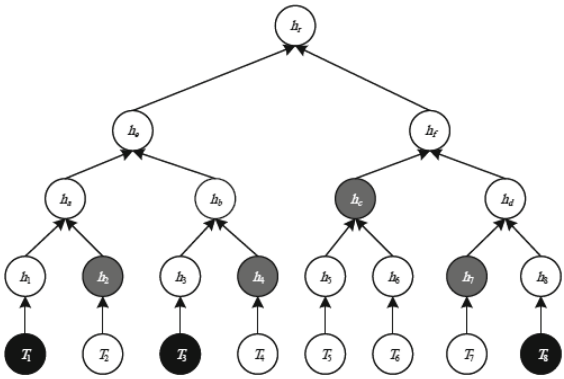

Fig. 6. The Cloud Storage Simulation Pattern

In this section, Python language is used to implement the scheme on a computer with Intel Core i5 CPU, Windows 10 system and 8 GB RAM. The back-end uses Node.js as the back-end operating environment, uses the Node.js framework based on IoC and also AOP to complete the function development, and adopts the architecture mode combining the single MVC mode and the micro-service, that is, the file service and the message can be integrated macroscopically.

Services are understood as two microservices relative to the entire system, and each specific service is a single application. Before outsourcing the data, we perform data preprocessing, including data chunking, generating MHT, and calculating homomorphic authentication tags, and then outsource the data to cloud services, in the Fig. 6, the cloud storage simulation pattern is demonstrated.

5 Conclusion

In this paper, we theoretically study the Artificial Intelligence (AI) assisted English linguistics teaching system and conduct the numerical simulation to validate the performance. Artificial intelligence can use the collected data to make the corresponding activities and feedback, provide adaptive and personalized learning and also personalized tutoring especially for the smart English guiding. The designed model contains the technology of face recognition, speech recognition and cloud storage. These models have significant implications for the understanding of how to use AI to assist the English linguistics teaching. The paper contributes the novel ideas on introducing the novel AI system, however, it also has certain limitations. Considerably more real-world teaching scenarios will need to be done to determine the validity and operability in the further study.

References

1. Ma, Yifang, Zhenyu Wang, Hong Yang, and Lin Yang. “Artificial intelligence applications in the development of autonomous vehicles: a survey.” IEEE/CAA Journal of Automatica Sinica 7, no. 2 (2020): 315–329.

2. Goralski, Margaret A., and Tay Keong Tan. “Artificial intelligence and sustainable development.” The International Journal of Management Education 18, no. 1 (2020): 100330.
3. Mak, Kit-Kay, and Mallikarjuna Rao Pichika. “Artificial intelligence in drug development: present status and future prospects.” Drug discovery today 24, no. 3 (2019): 773–780.
4. Vinuesa, Ricardo, Hossein Azizpour, Iolanda Leite, Madeline Balaam, Virginia Dignum, Sami Domisch, Anna Felländer, Simone Daniela Langhans, Max Tegmark, and Francesco Fuso Nerini. “The role of artificial intelligence in achieving the Sustainable Development Goals.” Nature communications 11, no. 1 (2020): 1–10.
5. Sun, Zhuomin, M. Anbarasan, and D. J. C. I. Praveen Kumar. “Design of online intelligent English teaching platform based on artificial intelligence techniques.” Computational Intelligence 37, no. 3 (2021): 1166–1180.
6. Guo, Meng. “Advantages and disadvantages of artificial intelligence in business English teaching.” In International Conference on Educational Innovation and Teaching Methodology, pp. 115–120. 2020.
7. Li, Ting, and Xuemei Zou. “Informatization of Constructive English Learning Platform Based on Artificial Intelligence Algorithm.” In 2022 International Conference on Frontiers of Artificial Intelligence and Machine Learning (FAIML), pp. 71–74. IEEE, 2022.
8. Batra, Gaurav, Zach Jacobson, Siddarth Madhav, Andrea Queirolo, and Nick Santhanam. “Artificial-intelligence hardware: New opportunities for semiconductor companies.” McKinsey and Company, January 2 (2019).
9. Jha, Sudan, Raghvendra Kumar, Mohamed Abdel-Basset, Ishaani Priyadarshini, Rohit Sharma, and Hoang Viet Long. “Deep learning approach for software maintainability metrics prediction.” Ieee Access 7 (2019): 61840–61855.
10. Roll, Ido, and Ruth Wylie. “Evolution and revolution in artificial intelligence in education.” International Journal of Artificial Intelligence in Education 26, no. 2 (2016): 582–599.
11. Gilakjani, Abbas Pourhossein. “Visual, auditory, kinaesthetic learning styles and their impacts on English language teaching.” Journal of studies in education 2, no. 1 (2012): 104–113.
12. Sekeroglu, Boran, Kamil Dimililer, and Kubra Tuncal. “Artificial Intelligence in Education: application in student performance evaluation.” Dilemas Contemporáneos: Educación, Política y Valores 7, no. 1 (2019).
13. Goldsmith, John A., and Geoffrey J. Huck. Ideology and linguistic theory: Noam Chomsky and the deep structure debates. Routledge, 2013.
14. Scontras, Gregory, Zuzanna Fuchs, and Maria Polinsky. “Heritage language and linguistic theory.” Frontiers in psychology (2015): 1545.
15. Alfaki, Ibrahim Mohamed. “University students’ English writing problems: Diagnosis and remedy.” International journal of English language teaching 3, no. 3 (2015): 40–52.
16. Liu, Jia, Alison Harris, and Nancy Kanwisher. “Stages of processing in face perception: an MEG study.” In Social Neuroscience, pp. 75–85. Psychology Press, 2013.
17. Sun, Yi, Xiaogang Wang, and Xiaoou Tang. “Deeply learned face representations are sparse, selective, and robust.” In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2892–2900. 2015.
18. Sun, Yi, Yuheng Chen, Xiaogang Wang, and Xiaoou Tang. “Deep learning face representation by joint identification-verification.” Advances in neural information processing systems 27 (2014).
19. Drago, Idilio, Marco Mellia, Maurizio M. Munafo, Anna Sperotto, Ramin Sadre, and Aiko Pras. “Inside dropbox: understanding personal cloud storage services.” In Proceedings of the 2012 internet measurement conference, pp. 481–494. 2012.
20. Rodríguez-Silva, Daniel A., Lilian Adkinson-Orellana, F. J. Gonz’lez-Castano, I. Armino-Franco, and D. Gonz’lez-Martinez. “Video surveillance based on cloud storage.” In 2012 IEEE fifth international conference on Cloud computing, pp. 991–992. IEEE, 2012.

21. Terry, Douglas B., Vijayan Prabhakaran, Ramakrishna Kotla, Mahesh Balakrishnan, Marcos K. Aguilera, and Hussam Abu-Libdeh. “Consistency-based service level agreements for cloud storage.” In Proceedings of the Twenty-Fourth ACM Symposium on Operating Systems Principles, pp. 309–324. 2013.
22. Slamanig, Daniel, and Christian Hanser. “On cloud storage and the cloud of clouds approach.” In 2012 International Conference for Internet Technology and Secured Transactions, pp. 649–655. IEEE, 2012.