# Design of Interactive Language Education Assistant System Based on Data Classification

Li Jiang(✉) and Xiaonan Yu

Nanchang Institute of Technology, Nanchang 330108, China
jl16023@163.com

**Abstract.** Aiming at the poor effect of language education, this paper puts forward the design method of language interactive education assistant system based on data classification, optimizes and improves the hardware structure of the system to ensure the operation quality of the system, further combines the principle of data classification, optimizes the function of the system software and simplifies the operation process of the system, Finally, experiments show that the language interactive education assistant system based on data classification can effectively break the limitations of space and time and improve the quality of language teaching.

**Keywords:** Data classification · Language interaction · Educational assistance

## 1 Introduction

With the development of education, computer technology, mobile Internet, and communication technology, computer teaching aid systems based on language interactive education aid platforms have been increasingly used in teaching, and have played an important role in the effective development of teaching activities [1]. With the rapid development of information technology and the continuous innovation of educational information construction, people have higher and higher requirements for teaching auxiliary systems. User needs are gradually developing in the direction of more functions, easier use, and faster efficiency. The role of network-based auxiliary teaching in modern education is becoming more and more obvious. The small remote system for the interactive education auxiliary system makes the communication between teachers and students, and between students and students more convenient and smooth, which is conducive to improving the learning effect [2]. The online auxiliary teaching system allows students to fully understand the teacher before class, preview the content of the course, easily obtain auxiliary materials after class, discuss and answer questions online, and also facilitate teachers to understand the student's dynamics in a timely manner, communicate and adjust the teaching method in a timely manner, So as to improve the quality of teaching and achieve better teaching effect.

Therefore, relevant researchers have done a lot of research on online assisted instruction system and achieved some results. Cao et al. Designed a method based on Net platform. The hardware structure of the system consists of user interface layer, service


S.-H. Wang and Y.-D. Zhang (Eds.): ICMTEL 2022, LNICST 446, pp. 183–195, 2022.
https://doi.org/10.1007/978-3-031-18123-8_14

selection layer and data management layer Teachers, students and other users enter their own identity information in the user interface layer, log in to the system, and click the corresponding program according to their own application needs. The business selection layer transmits the user selection instructions to the data management layer, and the data management layer selects the corresponding resources according to the user needs and feeds back to the user the interaction of the system is mainly reflected in interactive teaching and information interaction. Interactive teaching is reflected in online teaching between teachers and students; Information interaction is embodied in the information transmission of system information interaction model After testing, the designed system has strong pressure resistance, can respond to the application instructions of a large number of users in real time, and the interactive teaching effect is good, which improves the students' sense of self-efficacy. However, the operation of the system consumes a lot of energy and needs to be further improved. Hu proposed to design an online translation assistant system based on fusion of multilingual interaction. In the design of the system, the goal is to realize online learning. In order for learners to learn in the process of continuous correction, an end-to-end platform is designed. The platform integrates the machine translation server into the user interface most commonly used by professional translators, so that the machine can continuously learn from people's choices and adjust the model according to specific fields or user styles, Save the later editing work of learners and improve the auxiliary effect in the process of online translation However, the application of software in the system design is less, and there are some limitations.

Therefore, this paper designs an interactive language education assistant system based on data classification. The main design route of the system is:

(1) Based on the in-depth investigation of the background and current situation of the teaching assistant system, the hardware structure of the system is optimized.
(2) Combined with the principle of data classification, the structure of the course teaching assistant system is optimized. According to the teaching data of teachers and students' daily learning needs, analyze and design the teaching data management system which is more conducive to teachers' daily teaching work, and optimize the functional structure of the system.
(3) On this basis, the design system is divided into two parts according to the user type: student user data classification teaching information management, teacher data classification teaching information management and administrator user data classification teaching information management center to complete the system design.

## 2 Language Interactive Education Assistance System

### 2.1 Hardware Structure of Language Interactive Education Assistance System

Through the in-depth investigation of the background and current situation of the teaching assistant system, this paper summarizes some common basic needs of users, and analyzes the main user roles of the system. For example, in teaching practice, teachers and students play a major role, because language interactive educational assistance is a process of teaching and learning [3]. If a language interactive education assistant system

wants to assist teaching, it must provide certain data, tools, materials, etc. these resources are inseparable from the careful management of the administrator. At the same time, the administrator should also provide guarantee for the safety and normal order of the whole system. Therefore, the role of the administrator is also very important [4]. In order not to block the whole language interactive education auxiliary system, it is also necessary to consider non school personnel, that is, tourists. Their existence can share the excellent resources in the school with other scholars, continuously publicize the campus culture, and receive good feedback and suggestions from more talents and sages. Based on this, the hardware structure of the system is optimized as follows:

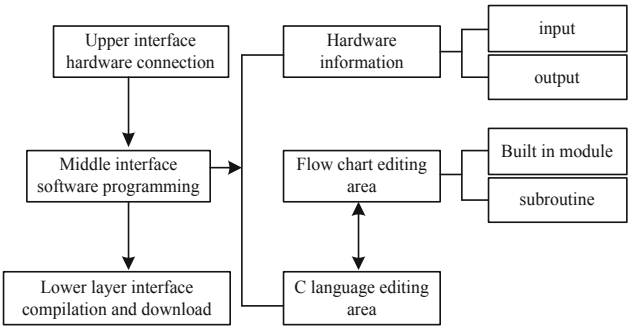


**Fig. 1.** System hardware configuration

In the system hardware configuration in Fig. 1, in order to realize the effectiveness of the auxiliary system, the hardware connection of the upper interface, the software programming of the intermediate interface and the compiling and downloading module of the lower interface are designed in this architecture, and the data output and input module is designed in the hardware. The PID switch and communication protocol of the network composite tracking sensor are designed in the system output module, as shown in the following Table 1 and Table 2:

**Table 1.** PID switch setting of composite tracking sensor

| Switch pin | Switch condition | Function |
|---|---|---|
| 2–5 | Open | Address value 0H01 |
| 2–5 | Shut | Address value 0H02 |
| 3–4 | Open | Find the white line and check that the white line light is on |
| 3–4 | Shut | Find the black line and check that the black line light is on |

In order to ensure the operation efficiency of the system, it is necessary to further optimize the main control board port and port connection information, and summarize the motherboard interface through a large number of practical search, as shown in the following list (Table 3):

**Table 2.** Composite tracing sensor communication protocol table

| Host command | C sensor return data | Effective value |
|---|---|---|
| P | Status values of 9 channels | 0–245 |
| PX1– PX9 | Monochrome data of corresponding 1–9 channels | 0–1352 |
| PX11–PX19 | Two color data of corresponding 1–9 channels | 0–1352 |
| PX21–PX29 | Threshold data of corresponding 1–9 channels | 0–1352 |
| PX31–PX39 | Current data of corresponding channels 1–9 | 0–1352 |

**Table 3.** Motherboard interface table

| Name | Motherboard number | Software identification | CPU port | Connecting parts |
|---|---|---|---|---|
| Analog interface | A3–A15 | P0–P9 | FD0–FD9 | Gray scale, infrared ranging, light intensity and other analog sensors |
| Digital interface | A18–A29 | F0–F8 | FB0–FB9 | Digital signal module |
| Electric drive interface | A28 | N0 | FB2–FB9 | On board dual motor drive output |
| Liquid crystal display | T6 | Y0 | – | On board LCD interface |

Combine the principle of data classification to optimize the system development process, build a language interactive education auxiliary information management platform, and improve the efficiency of the system. The specific structure is shown in the figure (Fig. 2).

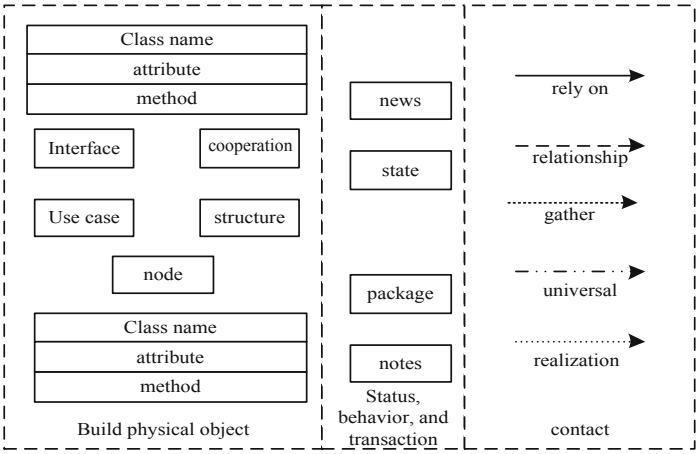


**Fig. 2.** Auxiliary information management platform for language interactive education

Based on the above method, the system hardware structure can be designed to better guarantee the quality of system operation and prolong the service life of the system [5, 6].

## 2.2 Optimization of System Software Function Structure

The system designed in this paper supports user types such as teachers and students. Users can query, modify and update their own information [7]. After students register, teachers can make an analysis and statistics on students' learning situation after each

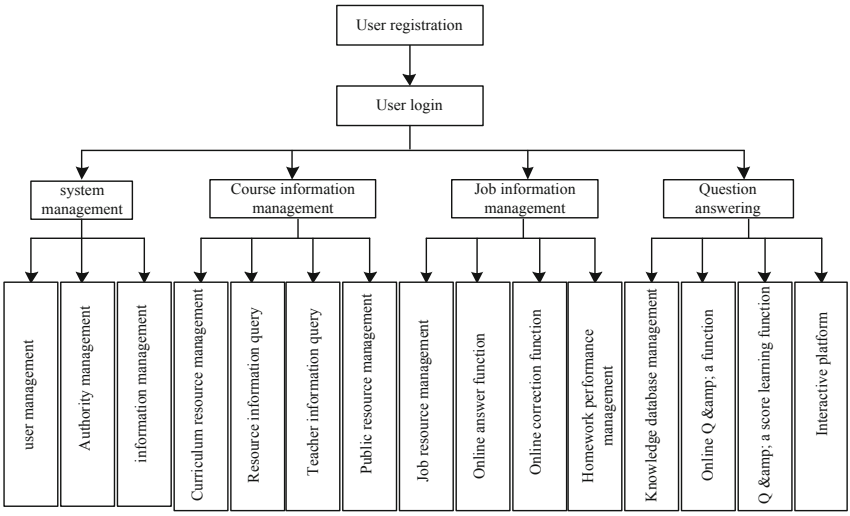


**Fig. 3.** System function structure diagram

class or after a period of time, timely understand students' learning dynamics, and feed back to students, so that teachers and students can interact and exchange course content in real time, so as to better improve the effect of course teaching [8]. Combined with the principle of data classification, the structure of course teaching assistant system is optimized. According to the work flow of teachers' curriculum teaching data and students' daily learning needs data, analyze and design a teaching data management system that can be more conducive to teachers' daily teaching work, and optimize the system functional structure, as follows (Fig. 3):

After analyzing the entire system of the system, it can be determined that this article will use the MVC architecture and methods, and use the B/S mode for auxiliary system development and design work. After the analysis, the specific architecture is included in this structure diagram as shown in the figure. Three-layer structure of presentation layer, logic layer and data layer (Fig. 4).

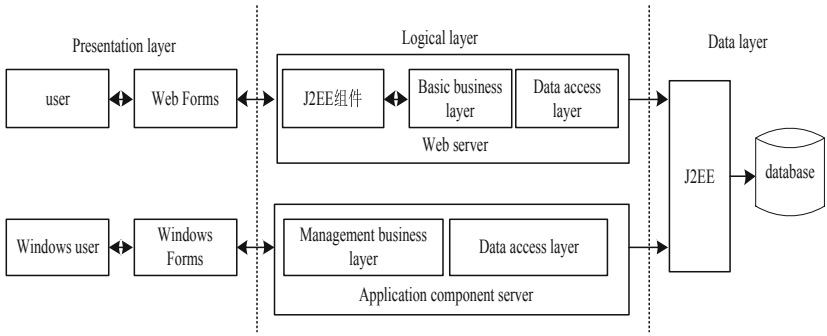


**Fig. 4.** Software architecture

Through the development and design of this system, it can be used by managers. At the same time, the development system and structure of the three-level architecture are also very clear, and the operation is relatively simple, which can greatly reduce the workload of the software package installation process [9]. In the system presentation layer, the main work is to use this interface to realize the conversion of data information. Commands can be input into the system in the form of machine language. This is very key for the presentation layer. Users can directly use this level. At the same time, the simplicity of the whole interface needs to be considered. The logic layer mainly operates and processes the data layer, issues operation commands through the presentation layer, and then processes the data [10]. After getting the data from the data layer, the data can be fed back to the system to improve the operation quality of the system.

### 2.3 Realization of Language Interactive Auxiliary Teaching

The designed system is divided into two parts according to the user type: data classification teaching information management for student users, data classification teaching information management for teachers and data classification teaching information management center for administrator users.

The system mainly provides students with six important modules: media learning environment, test question training environment, programming training environment, operation system, learning feedback and user information management. According to the actual language interactive auxiliary requirements, the system provides different user information management functions for users with different user permissions. Student users can only edit their own user information; Teacher users can edit their own user information and view some student user information, such as student number, name, class and other non private information; The system administrator can manage the user information of all users. There are two main management methods: manual editing and importing format files. Based on this, the system data information classification management process is optimized, as shown in the figure below (Fig. 5).

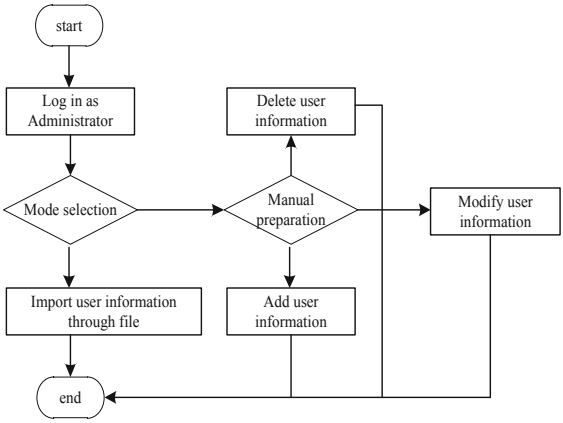


**Fig. 5.** System data information classification management process

The interactive assistance system provides multiple types of synchronous and asynchronous interactive tools, and combines them with the Steaching process of teachers and the learning process of students to form various "interactive toolbars" suitable for different teaching and learning stages. Facilitate convenient and efficient two-way communication and exchanges between teachers and knowledge transfer, between students and teachers, and between students and knowledge transfer. The main tools of the interactive auxiliary system The main tools used by the interactive auxiliary system are shown in the Table 4.

**Table 4.** Description of various interactive tools

| Type | Tool name | Feature description |
|---|---|---|
| synchronization | Hypermedia | The teaching content can be presented to students in various forms, such as text, graphics, animation and video, and connected with hypertext |
| | E-mail | With the communication means provided on the Internet, teachers can assign tasks and homework through e-mail, and answer one-to-one through e-mail |
| | Learning Forum | Online information service system can provide various resources, information and contacts to users in need |

Use the various teaching tools introduced above to build an interactive auxiliary subsystem to provide corresponding "interactive Toolbar" for different stages of network teaching and learning. The architecture model of interactive auxiliary subsystem is shown in figure (Fig. 6).

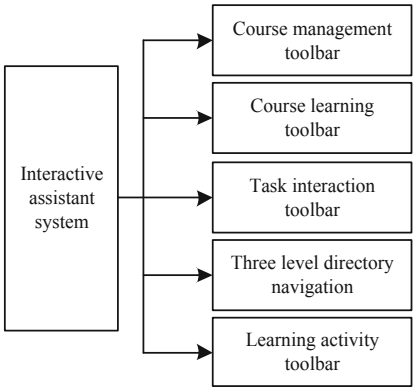


**Fig. 6.** Block diagram of interactive assistance system

In the process of data classification, a large number of data stored in the database do not meet the classification conditions. In order to enhance the effect of data classification, it is necessary to filter the original data. In order to ensure the horizontal and vertical inconsistency of each evaluation index, it is necessary to classify the data effectively. The classification of data is to consider the needs of users and recommend teaching resources to users who want to use them according to the needs of users. Therefore, this paper classifies the relevant data.

According to the characteristics of the teaching quality evaluation system, the data classification needs to meet the following rules 1: assuming that the college has $N$ total of a teachers, for convenience, it is advisable to set each teacher as $s_{kij}$ students, the number of all evaluation indicators is $t$, $k$ is the score of the j evaluation index of student

i, and $E_k(j)$ is the mean and variance of the j evaluation index of teacher K.

$$E_k\,(j)\;=\;\frac{1}{N}\sum_{N}^{i=1}s_{kij^{-t}} \tag{1}$$

$$D_k\,(j)\;=\;\frac{1}{N}\sum_{i=1}^{N}\left(s_{kij}-E_k(j)\right)^2 \tag{2}$$

If $D_k(j)=0$, it means that all students have the same data for $M$ evaluation indicators of teacher k, and the evaluation data will be deleted. $E_i$ is the mean and variance of student i is evaluation of teacher k:

$$E_i\;=\;\frac{1}{W}\sum_{W}^{k=1}E_i\,(k) \tag{3}$$

$$D_i\;=\;\frac{1}{WM}\sum_{W}^{k=1}\sum_{M}^{j=1}\left(s_{kij}-E_i\right)^2 \tag{4}$$

If $D_i=0$, it means that there is no difference in students' scores on all evaluation indicators *ot* of all teachers, then the data will not have a positive effect on data classification. Therefore, $\alpha$ and $\beta$ will not be selected. The mean and variance of students' teaching evaluation scores are as follows:

$$\alpha=E_i-E_k(j) \tag{5}$$

$$\beta=ot-D_i \tag{6}$$

Use overall classification accuracy $T_n$ and single classification accuracy $F_n$ to describe the performance of the SVM classifier, which is defined as follows: For the case where the data set sample distribution is relatively balanced, the overall classification accuracy evaluation index is used.

$$A=\frac{T_1+T_2+T_3+\ldots+T_n}{T_1+T_2+T_3+\ldots+T_n+F_1+F_2+F_3+\ldots+F_n} \tag{7}$$

For the unbalanced distribution of data samples, the single classification accuracy evaluation index is used.

$$C=T_i/(T_i+F_i) \tag{8}$$

In the formula, $T_i$ represents the number of correct classifications in the data sample of the type, and $F$ represents the number of incorrect classifications. When describing the performance of the classifier, the overall classification accuracy can objectively reflect the generalization performance of the SVM classifier, and the single classification accuracy can accurately reflect the pros and cons of the SVM classifier. Based on this, this paper divides the students' online behavior data samples into two categories under

the evaluation index with higher overall classification accuracy. Its functions include the formulation of various types of homework questions, the management of homework question banks, the release of homework, online submission, online correction, and the statistical analysis of homework scores and score information query functions. Program function introduction: Multiple choice questions and judgment questions can be automatically corrected and the scores displayed after the homework is submitted. Students can view the homework history and correct answers from the homework record. Teachers can manage homework question types and question banks, make online corrections to some question types, and give scores and comments, as well as statistical analysis and management of results. Work flow: After the students log in to the system, they will complete the homework according to the teacher's assignment.

Students ask questions to the system in the way of natural language, and the intelligent question answering agent gives answers. Intelligent Q & A agent has the ability to learn, accumulate knowledge and improve the ability to answer students' questions. After users ask questions, the system analyzes the keywords, matches them with the answers, and feeds back the results. Both teachers and students can score the answers given by the system, so as to continuously train the intelligent question answering agent of the system.

The steps are as follows (1) The user enters the question in natural language, and the intelligent answering agent obtains the content of the question and displays it in the dialog box. (2) The intelligent answering agent analyzes the question, gets a number of keywords, and filters out the redundant words by merging synonyms, etc., to get the keywords that are actually needed. (3) According to the obtained keywords, query in the knowledge database to get the most consistent answer. (4) Return and display the answer. Based on this, the effective design of the system is realized, and the effect of auxiliary teaching is effectively guaranteed.

## 3 Analysis of Experimental Results

Any set of system programming software needs supporting software, which is developed using object-oriented V programming software. On the basis of this development software, the supporting software is used for development, and the system uses programming management software to support the software, and the system is programmed by flowchart programming software. And C language programming software is composed of two parts, programming management software is used to establish the system low-level module software and set up the program running environment. The flowchart programming software and C language programming software rely on AVR GCC support, and need to install Winavr GCC at the same time. This software has established environment settings. The default GC version used is: Winavr20070525. The built-in software module is the part executed in the flowchart programming language. This part tests whether its functions can be dragged and used correctly. The test results are shown in the Table 5.

After the development and design of each functional module of the system, it is necessary to fully test the function and performance of the system, so as to prevent abnormalities in the process of system operation. At present, the main test methods

**Table 5.** Built in software module test table

| Type | Can I use it | Operation | C language conversion status |
|---|---|---|---|
| Conditional judgment | Sure | Stable | Accuracy |
| Count repeat | Sure | Stable | Accuracy |
| Condition repetition | Sure | Stable | Accuracy |
| Do repeat | Sure | Stable | Accuracy |
| Repeat forever | Sure | Stable | Accuracy |
| Assignment statement | Sure | Stable | Accuracy |
| Toms delay | Sure | Stable | Accuracy |
| MS Delay | Sure | Stable | Accuracy |
| Sensor | Sure | Stable | Accuracy |
| start-up | Sure | Stable | Accuracy |
| stop it | Sure | Stable | Accuracy |

are black box test and white box test. Black box test is to detect and analyze each functional module of the system to ensure that the functional modules of the system can operate normally. There are some differences between white box test and black box test, which requires designers to have a certain understanding of the structure of the program, and carry out internal detection and analysis. When testing the system function, first of all, it is necessary to test the data interface of the system, determine the relevant parameters and information input by the user in the test process, and also meet the following requirements: (1) divide and test the functional modules of the system, complete the test analysis of the main functional modules, and complete the big data test. (2) When testing the network of the system, users need to log in to the system many times, so as to determine the user's actual operation, which can ensure the normal use of the system. During the specific test, the personnel participating in the test need to adjust according to the parameters, as shown in the Table 6:

Teaching assistant management includes various functions such as examination management, score management, and student attendance management. For the functions of the system, these functions are the main functional modules in the entire system, as shown in the figure (Fig. 7).

The test of the system is not only considered from the function, but also from the performance and other aspects. In particular, different users will participate in the test process and need to be tested and processed according to different user roles. In the process of system test, all functional modules are closely connected with the test process. In addition, the process is easy to understand, the user interface is very friendly and the operation is relatively simple. After the function test, the abnormal performance test is also required. The specific method is that the user inputs the abnormal value to judge whether the system can avoid the occurrence of abnormal conditions during operation. At the same time, the system can prompt that the value entered by the user is illegal. In the test results, it can be determined that there are 160 use case designs selected

**Table 6.** Functional test cases

| Control | Expected results | Is it stable |
|---|---|---|
| Software main interface and menu | Operate the main page, including registration, login, use, etc., to determine whether the main interface and menu functions of the software are stable | Stable■ unstable□ |
| Message publishing management | Administrators use this function to publish message data | Stable■ unstable□ |
| Binding effect of official documents | Staff can manage official document data | Stable■ unstable□ |
| Teaching resource management | The system can sort out teaching resources | Stable■ unstable□ |
| Auxiliary teaching | The system can provide courseware, audio and video for auxiliary teaching | Stable■ unstable□ |
| system maintenance | The administrator can back up the database and download the database | Stable■ unstable□ |

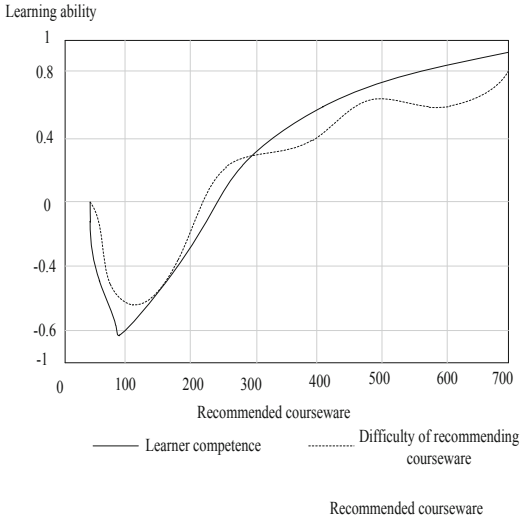


**Fig. 7.** Evaluation of the recommended effect of courseware

this time, The total number of functional modules involved is 400, and there are 380 normal test items, accounting for 95% of the test proportion; 15 items with light defects, accounting for 3.75% of the test proportion; There are 5 moderate defects, accounting for 1.25° of the test proportion. Although the system runs well, there are still some errors, mainly as follows: users of any identity can view all functions when entering the system. Although users without permission cannot use these functions of the system,

there are still defects. In addition, the description of the interface is inaccurate. These test problems have been corrected after the test results, and all have been changed and improved before deployment.

## 4 Conclusion

This paper designs a language interactive course assistant system based on data classification. The system design includes hardware and software. In the hardware design, the overall optimization of the system module is optimized, and the connection layer and other modules of the system are designed. In the system software design, the data classification method is used to classify the extracted user demand data to improve the performance of the auxiliary system. The experimental results show that the system designed in this paper has good performance. However, in the current research, more user data and user needs should be considered in the system design, which will be helpful in the future.

## References

1. Tassi, A., Gigante, D., Modica, G., et al.: Pixel- vs. object-based landsat 8 data classification in google earth engine using random forest: the case study of Maiella national park. Remote Sens. **13**(12), 2299 (2021)
2. Castrogiovanni, P., Fadda, E., Perboli, G., et al.: Smartphone data classification technique for detecting the usage of public or private transportation modes. IEEE Access, **99**, 1–1. [1] (2020)
3. Li, G., Ouyang, S., Yang, Y.: A study on the construction of a culture pedagogical network learning space – the CASH curriculum idea. Int. J. Emerg. Technol. Learn. (iJET) **14**(17), 73 (2019)
4. Yang, K., Ma, Q.: To improve the quality of distance education and the simulation of effective resources under the big data. Comput. Simul. **34**(4), 212–215, 334 (2017)
5. Omar, S., Hanifah, M.A., Rathakrishnan, L., et al.: The use of kahoot interactive quiz in the teaching of Malaysian nationhood studies course at universiti Utara Malaysia. Solid State Technol. **63**(3), 177 (2020)
6. Kassabolat, A., Kadirsizova, S., Kozybayeva, M., et al.: Future teachers' opinions on preparation and use of interactive materials in teaching. Int. J. Emerg. Technol. Learn. (iJET) **15**(23), 121 (2020)
7. Muoz, L., Villarreal, V., Morales, I., et al.: Developing an interactive environment through the teaching of mathematics with small robots. Sensors **20**(7), 1935 (2020)
8. Sprute, D., Tnnies, K.D., Koenig, M.: A study on different user interfaces for teaching virtual borders to mobile robots. Int. J. Soc. Robot. **11**(3), 1–16 (2019)
9. Boschmans, S.A.: Becoming a pharmacy assistant and technician: education and training. Encyclopedia Pharm. Pract. Clin. Pharm. **15**(02), 15–24 (2019)
10. Liu, S., Fu, W., He, L., et al.: Distribution of primary additional errors in fractal encoding method. Multimedia Tools Appl. **76**(4), 5787–5802 (2014). https://doi.org/10.1007/s11042-014-2408-1