Journal of Engineering Education Transformations ,
Volume 37 , No. 2 , October 2023 , ISSN 2349-2473, eISSN 2394-1707

10.16920/jeet/2023/v37i2/23159

The Utilization of Artificial Intelligence Based Chatbot in Interactive Learning Media

Baharuddin[1], Muhammad Dominique Mendoza[2], Olnes Yosefa Hutajulu[3], Hesti Fibriasari[4]
[1 2 3] Faculty of Engineering, Universitas Negeri Medan
[4]Faculty of Language and Arts, Universitas Negeri Medan
[1]baharuddin.bah@gmail.com
[2]aenaen@unimed.ac.id
[3]olnes.hutajulu@unimed.ac.id
[4]hestifibriasari@ymail.com

Abstract: After the spread of the COVID-19 virus, the world of learning is no longer the same as it was before the pandemic, especially at the university level, which requires students to be more active in finding lots of learning references via the internet. Active learning has a close relationship with the use of artificial intelligence technology, in this study researchers designed an interactive learning media that applied Artificial Intelligence (AI) Assisted Chatbot. The multimedia development method used in this research is the Multimedia Development Lifecycle (MDLC) method. Interactive learning media is built using Android Studio, and the retrieval-based chatbot is built with Python, Tensorflow, Hard (Extension for Tensorflow), NumPy, and Matplotlib. The researcher applies the White box and Black box testing methods to test whether the learning media that have been made have worked according to the user's needs. The chatbot model that has been applied to this learning media when tested with the BLEU metric gets a result of 0.1117, this shows the chatbot produces good answers and provides credible learning references. There is a difference in student learning outcomes after using chatbot-based learning media in basic electronics subjects, the utilization of chatbot-based learning media is able to improve student learning outcomes, especially in the student weaknesses section, the chatbot can provide detailed explanations and guide students in solving linear equation problems.

Keywords: Artificial Intelligence; Chatbot; Interactive Learning Media; Retreival-based Chatbot;

Baharuddin
Faculty of Engineering, Universitas Negeri Medan
baharuddin.bah@gmail.com

1. Introduction

After the spread of the COVID-19 virus, the world of learning is no longer the same as it was before the pandemic, especially at the university level, which requires students to be more active in finding lots of learning references via the internet [1]. Active learning requires students to study independently so that meaningful learning can be created in accordance with the learning outcomes that have been set at the beginning of the lecture. Active learning has a close relationship with the use of artificial intelligence technology, in this study researchers designed an interactive learning media that applied Artificial Intelligence (AI) Assisted Chatbot [2]. All people involved in the world of education try to make their work with computers easier, they use various text or voice interfaces, with which they enter many commands into the computer and then respond to them. The majority of educational studies of artificial intelligence have concentrated on software or system development [3][4]. AI is also used as the basis for development, such as AI-assisted decision-making design [5], conversational agent design [6], and AI algorithm development [7]. The vast majority of research on learner technology acceptance has concentrated on technical knowledge, also known as computer self-efficacy or digital literacy rather than subject domain knowledge [8-12]. The purpose of this research is the implementation of a chatbot (a text interface where the computer reacts to text input by responding to its output) into an interactive learning media. The essential criteria for the developed chatbot are that it delivers good replies and trustworthy learning references, and that it responds clearly and without numerous grammatical errors. Previous study has demonstrated the relevance of chatbots in swiftly providing acceptable replies in response to enquiries, although the answers supplied are frequently inadequate [13]. According to the findings of 80 studies on chatbots and their applications in education, there are numerous types of educational chatbots now in use that influence student learning or improve services in various domains [14], therefore a chatbot must provide credible learning references. Based on prior study on chatbots that concentrated on grammatical errors in chatbots, the need to improve chatbot performance bodes well for future improvements, particularly in the section on grammatical errors [15].

This research explores the gap by collecting data on student and teacher perceptions of the use of chatbots in interactive learning media. Although chatbots can help improve interaction between students and learning materials, there has not been much information on how chatbots are received by students and teachers as a learning medium [16][17]. Creating chatbots for interactive learning media requires expertise in interface design, AI development, and an understanding of how students learn. However, not much research has explored the gaps in the development of chatbots suitable for interactive learning [18]. Although AI technology is growing, there are still limitations in the use of chatbots for interactive learning [19]. Some of these limitations are difficulty in processing natural language and difficulty in providing appropriate feedback. Therefore, this research will explore the use of AI chatbots in interactive learning and find solutions to overcome some of the shortcomings. This research will help improve the understanding of how chatbots can be effectively used in interactive learning and contribute to the development of AI technology in the context of learning.

The contribution of this research is to introduce a new solution in overcoming the problems experienced by students. AI-based chatbot as a solution to improve learning quality. This solution has not been widely applied in the field of education, this research is expected to provide an original contribution in the use of chatbots for learning purposes. The use of AI-based chatbot in interactive learning media can improve learning effectiveness. This will benefit students who can understand difficult concepts more easily. In learning Basic Electronics there are often concepts that are difficult for students to understand, especially about linear equations. By using AI-based chatbot in interactive learning media, students can easily understand these concepts. The chatbot can provide more detailed explanations and provide exercises to deepen students' understanding. Thus, students can understand these concepts more easily and effectively, so that their academic performance can improve. This research can provide practical benefits for students in improving learning effectiveness. In addition, the use of chatbots in learning can also reduce the burden on teachers in providing individualized explanations to each student. The chatbot developed in this research is specifically designed for learning needs. This chatbot not only provides explanations and exercises, but can also customize learning to student needs and provide instant feedback. The chatbot developed in this study uses AI technology to provide fast and accurate responses to student input. The AI technology used in the chatbot can accelerate the learning process and help students overcome learning difficulties. The chatbot developed in this study can collect data about its use, such as feedback from students, chatbot performance, and student learning outcomes. This data can be analyzed to improve and enhance the performance of the chatbot in providing more effective learning.

Chatbot testing is carried out in 2 stages, the first stage is testing the chatbot response according to the evaluation algorithm (White box testing), and the second is user testing. The students and lecturers were given a learning media that had a chatbot feature and had a conversation with it, after which they were given a questionnaire with questions about the quality of the conversation with the chatbot. The entire program is written in the Python programming language and uses the Tensor library stream for mathematical calculations. The implemented chatbot only functions for communication between the computer and the user, this chatbot is not intended to completely replace the lecturer. The chatbot also provides useful information input to students for learning references in fields that they are less capable of and provides input information to lecturers about mapping data on student evaluation levels in the class.

2. Review of Related Literature

A. Chatbot

The ELIZA program developed by the German professor of mathematics Joseph Weizenbaum, which in 1966 simulated a conversation between a psychologist and a patient (user), is considered the first implementation of an intelligent dialogue system [20]. Keywords are searched for in the entered text, to which a priority number is assigned according to their importance - the so-called rank. Each keyword is associated with a rule according to which the input sentence is transformed. A big advantage of ELIZA is the use of a script that is not part of the program itself, but functions as external data containing keywords with appropriate transformation rules. This makes extension to another natural language trivial. All you have to do is rewrite the script in the language of your choice. The problem may occur if the entered sentence does not contain any keyword. At this point, the system will write down a pre-prepared neutral sentence, such as "Please continue," "That's very interesting," or "Tell me more about it." In ELIZA's case, it's not about real intelligence, it's just about using tricks, and thanks which give the system a trustworthy impression. But despite this fact, many people were convinced that they were communicating with a real psychotherapist and not with a computer.

Based on their principle of operation, we divide chatbots into two categories. Firstly, a chatbot is based on a set of simple rules that transform a given input into a suitable output (the output can be a declarative or interrogative sentence). Therefore, it can only respond to inputs defined in the rules. It cannot parse other inputs. This group includes ELIZA and systems with a specific focus - for example, user support (search for travel connections, for gram cinemas or theaters), these are so-called retrieval-based chatbots [21]. The second group includes advanced implementations using neural networks. They understand natural language not only by pre-defined commands but improve their skills with each conversation, these chatbots are called generative chatbots. The implementation of chatbots from this group is much more demanding than the implementation from the first category.

Furthermore, these programs can be divided on the basis of obtaining information. Some systems use their own knowledge database, which is fixed or automatically expanded during communication with the user. The second option is to search the web - this uses, for example, the PAR system to automatically answer questions. The domain indicates which inputs the chatbot is able to respond to and which, on the contrary, it cannot or can only partially respond to. That's why we divide domains into two basic types: Open-domain, with this type of domain, there is no limit on where the input sentence for the chatbot comes from. There is no need to have a pre-defined purpose for the conversation. Thus, the user can have a conversation on any topic and does not have to stick to any particular topic. Unfortunately, with an infinite number of topics, it also increases the difficulty of creating a chatbot that could respond correctly to everything needed. Closed Domain, a closed domain chatbot can only respond to certain inputs from the user. Both input and output are limited by the fact that the chatbot is trying to achieve a specific goal [22]. As a result, such a system cannot respond to all possible input cases. Systems that mostly use a closed domain are, for example, technical support or assistants helping, for example, with shopping. This is due to the fact that we do not require these systems to communicate on various topics, but they only have to properly help the user with a specific request.

B. Retrieval Based Chatbot

Retrieval-based chatbots or also called rule-based chatbots are much simpler to design and implement [23]. This is mainly due to the fact that it uses a specific database where there are already predefined questions and answers to these questions. A chatbot can contain several different databases. The output of the chatbot model that is implemented as retrieval based is a score that shows the accuracy of the answers to user questions. The answer is then selected from the set of sentences in the database that has the highest score. For simpler types of chatbots, the sentence database can be customized. This means that chatbots have sentences stored in this database before they even start using them. With more complex chatbots, sentences are stored in the database even while the chatbot is in use. Expansion is performed according to new user input, and this feature may also result in inappropriate data being stored in the database. This happens, for example, when a user gives a nonsensical answer. There are also types of chatbots that do not contain a database of sentences and answers. This type of chatbot then search the Internet for answers to questions entered by users [23][24].

Various evaluation metrics are used to find out how accurate the answers are given by a chatbot built on a retrieval-based implementation. One of the most frequently used metrics is the metric called recall@k. Recall@k selects the best answers from 10 possible (predetermined) answers [25]. One of these ten answers is correct and the others are so-called distractors. If the correct answer is among the selected ones, then the test is evaluated as successful. If we set k to 10, the probability of correct selection is 100%. It follows that the smaller k, the lower the probability of correct selection, and vice versa.

C. Generative Chatbot

Generative models have no predefined answers, they generate new answers [26]. They are mostly based on the same principle that models for translation from one language to another work on. In this case, however, there is no translation between languages, but a transformation of one input into its given output (answer to a query). The ideal way to measure the accuracy of a chatbot's responses is through human factor judgment. Such a metric is the BLEU algorithm, which is mainly used to determine the quality of the translation for machine translations of texts into other languages [27]. This method can also be used to determine the quality of a newly generated answer for a generative chatbot. The fact remains that research shows that there are no known metrics used, which could evaluate the answers as well as human judgment.

Both types of chatbots have their advantages, but also disadvantages. Retrieval-based chatbots do not make grammatical errors. The error could only occur if it was contained in an already predefined response in the database. But on the other hand, they are not able to handle being forced to respond to cases for which they do not have this predefined response, nor can they respond to information that has already been mentioned in a previous conversation - names, for example. Generative models are more useful in this regard. They can refer to the information mentioned in a previous conversation and can also give the impression that the user is communicating with a person. But these models are hard to train. Unlike retrieval-based models, they can respond with grammatical errors, mainly for long sentences. Another disadvantage is the need for a large amount of training data.

D. Topology

This section deals with the correct choice of neural network, the simplest type is a feedforward neural network. Here, the output of all neurons from the lower layer is fed to the input of the higher layer. The output layer, it is the output data of the feedforward neural network itself. This also results in the very name of this type of neural network - feedforward neural network. A multilayer neural network is capable of solving even non-linear problems, it is capable of approximating complex functions and serves as a general approximation model [28]. We must be careful that the number of layers is not proportional to the improvement of the results. To solve many problems, it is convenient to use the architecture of one input, process, and output layer. To get an idea of what it looks like, please look at Fig. 1.

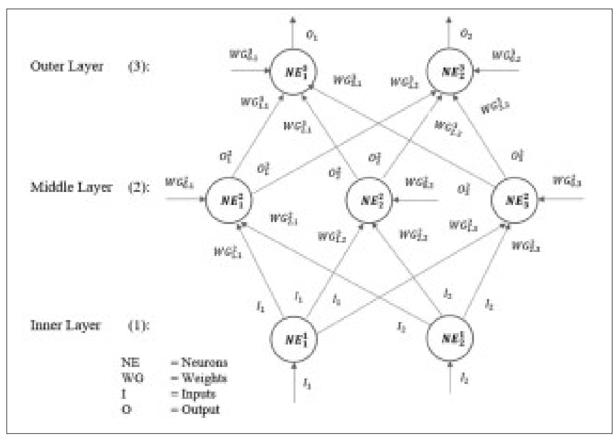

Fig. 1: Marking of neurons, weights, inputs and outputs in a neural network

Furthermore, it is necessary to choose an appropriate number of neurons for the correct generalization of the problem during training, and it is also necessary to determine what data is involved in training. It is practically possible to start with a small number of neurons and further increase it so that it is able to approximate the function. It is also possible to use genetic algorithms to find the correct (suitable) number of neurons. We need to know that it will be a large neural network. If the number of neurons is too small, the neural network is not able to generalize properly, but on the other hand, if we choose too large a number of neurons, overtraining may occur. Overtraining of a neural network is an event in which the neural network shows very good results on the training data, but when used in real life on the test data, the results are poor quality, cannot be used [29]. Unfortunately, overtraining is very common and there are techniques for how to prevent this problem.

E. Interactive Learning Media

Interactive means influencing each other where there is a reciprocal relationship between the user and the program where the user responds to the request / program display, then the program presents the desired information [30]. The interaction that students make through the buttons available on the program can provide a hands-on learning experience. Interactive multimedia can be controlled by its users so as to provide a direct experience for students. In connection with this, the development of interactive multimedia that will be carried out is by providing a chatbot assistant in learning media so that students can gain direct learning experiences by self-study and students feel they have close friends as a place to exchange ideas. Interactive multimedia using chatbots is an innovation in technology that can be adapted to students' learning needs [31]. This interactive multimedia is an innovation that is needed as a learning medium that is in accordance with advances in technology and information. As a step taken by researchers to realize this, researchers developed interactive learning multimedia assisted by Chatbots. Students will be motivated if they are interested in the material being studied, therefore the material must be packaged as attractively as possible with easy-to-understand explanations. This can be interpreted that the development of interactive multimedia that will be carried out needs to pay attention to the presentation of the material so that it looks interesting and in accordance with the characteristics of students [32]. The originality of this research is that this research focuses on the development of interactive learning media applied to Basic Electronics course learning and provides significant benefits for students. This research focuses on the needs of users, especially students, and how interactive learning media using AI-based chatbots can help meet these needs. This learning media with chatbots provides solutions to students who need learning media that can help them learn more effectively and efficiently. The use of AI-based chatbots in interactive learning media can help students understand difficult concepts and provide solutions to problems faced by students quickly and easily. Thus, the use of AI-based chatbots can help meet the needs of students in improving the quality of learning. One example of a difficult concept in Basic Electronics is linear equations. Many students have difficulty in understanding this concept because there are many formulas and equations that must be remembered. By using an AI-based chatbot in interactive learning media, students can ask questions and request explanations about the concept of quadratic equations. The chatbot can provide detailed explanations and guide students in solving linear equation problems. In this case, the use of AI-based chatbot can help students in understanding difficult concepts such as linear equations and improve the quality of learning.

3. Methodology

A. Learning Media Development

The multimedia development method used in this research is the Multimedia Development Lifecycle

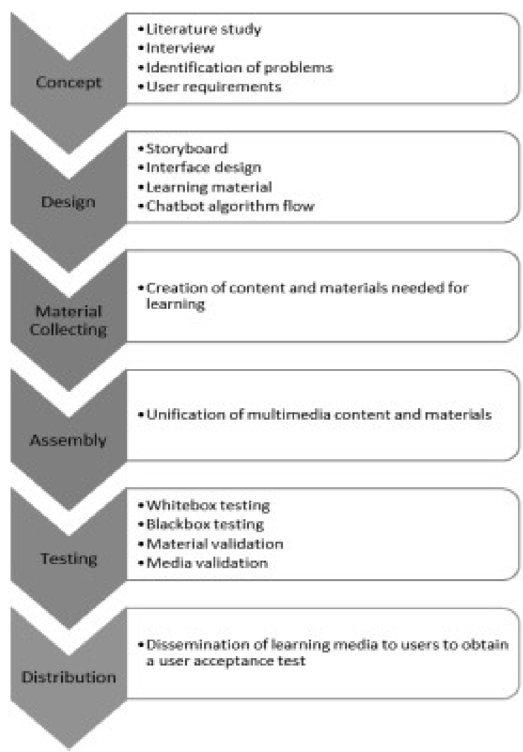

Fig. 2: Learning media development

(MDLC) method (Fig. 2). This method is used for developing interactive multimedia applications and integrating Chatbot into the learning media that has been created. In this method, the main focus in the development process is the aspect of functionality and content [33]. The researcher developed a learning management system (LMS) focusing on a chatbot that can select the additional materials needed by students to study each lecture topic that best suits the needs of users and produce feedback that can be understood by students. The chatbot will also display materials in various teaching materials that have been previously prepared on the LMS.

B. Concept Stage

At this stage, identification of problems is carried out by studying literature obtained from books, journals, and other references related to this title, then conducting interviews with prospective lecturers and students using learning media, then analyzing existing conditions related to learning methods. used by students and teachers. After that, identification of needs is carried out, both application needs and user needs, and determines the purpose of the application, then identification of subject matter is carried out. Interactive learning media built using Android Studio (used to create learning media based on Android), Python (a high-level scripting language, it offers dynamic checking of data types and supports various programming paradigms, including object-oriented, imperative, procedural, or functional), Tensorflow (an open-source library for mathematical calculations using graphs for flow), Hard (Extension for Tensorflow, it simplifies the definition of neural network models), NumPy (a library for complex calculations, to enable efficient work with N-dimensional), Matplotlib (a data plotting library, to generate curves, histograms, spectra, and tables) [34] [35]. This application is made with the concept of interactive multimedia so that someone who wants to learn can choose the material they want to learn, in delivering the material it will be packaged as attractive as possible and not boring with a chatbot assistant. The output of this stage is to define the application objectives, system requirements, materials lessons, and storylines.

C. Design and Material Collecting Stage

At this stage, the research team designed the interface, made a storyboard, and the chatbot algorithm flow. The output of this stage is the creation of learning storyboards, system navigation structures, and evaluation plans. In the Material Collecting stage, researchers collect application supporting elements such as images, audio, music, and others. The outputs of this stage are pictures, music, animation, audio, and books that are suitable for children who have a visual learning style.

D. Assembly Stage

The assembly stage is the stage where all multimedia objects are created based on the storyboard and navigation structure that comes from the design stage. In this stage, illustrations, audio and video are made, as well as programming, this work is done by entering the contents of the material into the screen as contained in the design. The output of this stage is to determine hardware requirements specifications, software requirements specifications, and application interfaces.

E. Testing Stage

This stage is carried out after the creation stage is complete and all data has been entered. First, White box testing is carried out to ensure that the program flow, especially the Chatbot algorithm, has the appropriate response as expected and provides the right information feedback [36]. Then black box testing is carried out, this test is carried out to check whether the application functionality is running well or not and identify errors related to software functionality errors that appear in the output error. The test results for functionality are successful for all existing functionality and are in accordance with the expected output. Then the media expert validation test and material expert validation were carried out. The next stage is the user acceptance test, which is a trial for system evaluation carried out by users, namely students and lecturers with the aim of obtaining approval for the system being tested and ready to be used [37]. The results are known by asking the user to answer the contents of the questionnaire that the research team asked the user to measure the quality of the display, application interaction, presentation of material, and user interaction, in accordance with the wishes of the user.

F. Distribution Stage

This stage is carried out if testing has been carried out and the learning media is declared feasible. The finished application is distributed by uploading the application to a Google Drive link, where later this application will be downloaded by students online.

4. Results and Discussion

A. Concept Stage Result

After conducting direct interviews with the subject lecturers and students involved, the researchers made direct observations and attended the class and got information that learning is still less effective and still one-way between lecturers and students, so we need a virtual assistant who can be trusted with information and can help students in completing the subject. Researchers also interviewed students to get feedback and obstacles during learning. Based on this, it is conceptualized to create interactive learning media containing subject matter, evaluation questions, and a virtual assistant using a chatbot.

B. Design Stage Result

The researcher uses the storyboard application to design the learning media interface, then implements it into Android Studio. The researcher created a script Start_chatbot.py which aims to process the arguments and then call the required functions. One of the most important functions of the chatbot is located here, and that is the modelCreation() function. As the name suggests, this is a feature for creating a chatbot model. More precisely, it is not just about creating one model, the user can choose to create one of three models. In this script, there is a ChatManager class that manages the chat between the user and the chatbot. The getAnswer() function is called to get the chatbot's answer. It is called from another function found in this script, which is the chat() function. This function controls the response dump chatbot to the user's terminal. Another script that controls the output of the chatbot's responses via the GUI, not to the terminal, is the gui.py script. Here we can find the GUI class that is used to create a chat GUI for the user. It also uses the getAnswer() function to list the answers, as was the case with the list to the terminal.

To start the chatbot, the function start_chatbot.py must be called. The chatbot can be run in three modes – training, chatting, and testing. If we run this script with the -t parameter or with the –train parameter, the train() function will be called to start training. This function needs three parameters. The number of the model on which the neural network will be trained, the training data limit, and a boolean parameter. The last parameter tells whether to continue training from the last stored training weights or whether to train from the beginning.

Another mode that can be used to start the chatbot is -c or –chat. This will start the chat mode with the user. Here the chat() function is called, which needs two parameters - the model number and the boolean value of GUI. The model number indicates which model will be used to chat with the user, the GUI parameter then determines whether the chat will take place only in the terminal (set to false) or if set to true, the chat will take place through the UI. In order to test the chatbot, it is necessary to run the script using the –test the chatbot will start testing, and the testBLEUFromLastSavedWeights() function is called, which requires entering three parameters: the number of the model to be tested, the limit for testing,

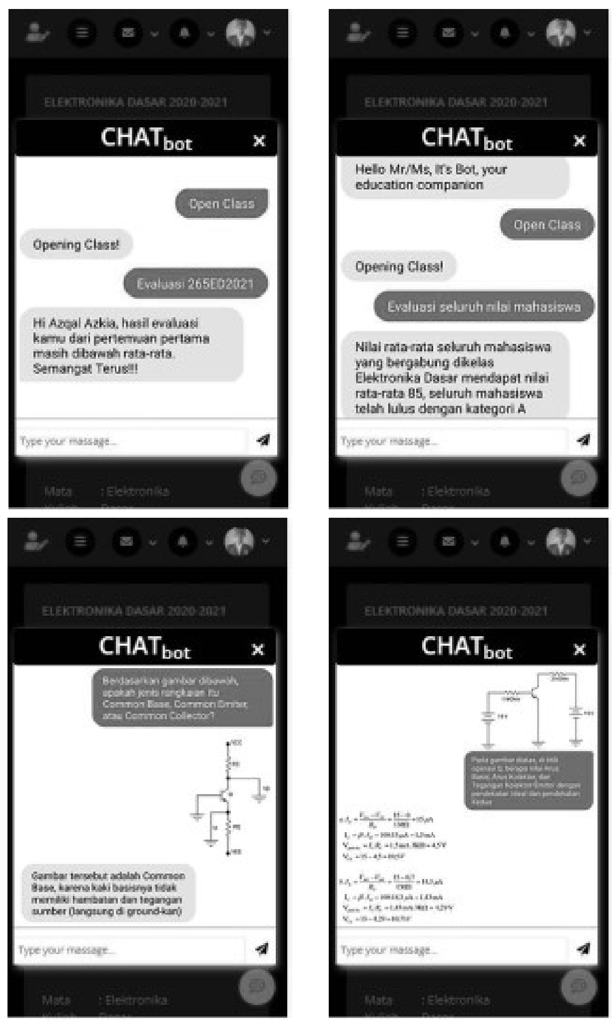

Fig. 3 : Chatbot interface

and last one is a boolean value that indicates whether testing of all models or just one specific one was specified. If no parameters are entered when starting the chatbot, the necessary score for the start will be set automatically. These implicit values are the determination of the mode in which the chatbot should work, and the model that will be used - this will be the basic (first) model. Furthermore, all boolean values will be set to "False", so the chatbot testing will not take place, and communication with it will be via the terminal. Furthermore, other parameters can be used, not only parameters for determining the mode of the chatbot. For example, it is possible to use the -g or –GUI parameter, which is a boolean value that specifies in which environment the user will interact with the chatbot. If the value is set to "True", the GUI created for communication is opened. Another possible parameter is the parameter – model=[model number] intended for setting the model to be used in the given mode. If the user enters a number other than the number of models, the program is terminated with the error "Wrong model number. Select number 1-3". The last parameter that can be specified when starting the chatbot is the -h or – help parameter. After entering it, a "help message" will be displayed and the program will end. This message is intended to inform the user, and there is a description of the chatbot as well as help on how to start the program.

C. Material Collecting Stage Result

At this stage, the researcher collects materials in the form of animated assets, learning videos, audio effects, and collects lecture materials that will be displayed in learning media. Then the researcher also updated the chatbot language database using everyday language used by students and lecturers to minimize the occurrence of words that were not understood by the chatbot and minimize bug and errors.

D. Assembly Stage Result

At this stage, all the features and content that have been created will be checked again to ensure that all the material and content displayed is in accordance with the storyboard that was created in the previous stage. The results of this stage are illustrated in the use case diagram in Figure 4 below.

The learning media model created by the researcher can be seen through the picture above, students can access the material menu, evaluation, chatbot, learning outcomes, and guidelines for using

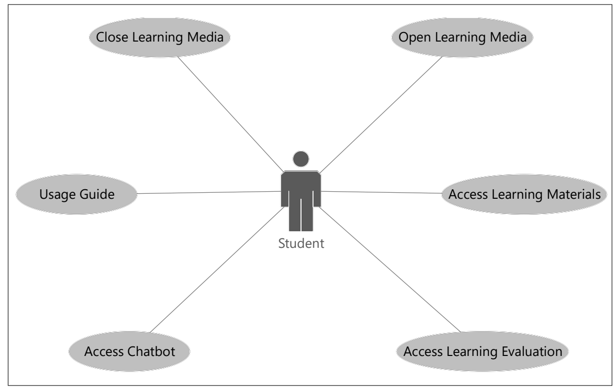

Fig. 4: Students' usecase diagram

learning media. In the material menu, students can see case-based learning materials and consist of 14 meetings with different topics at each meeting. On the student evaluation menu, case-based questions are presented for each learning topic. In the learning outcomes menu, students can see a summary of the assessments of all learning activities that have been carried out by students.

The chatbot menu acts as a student virtual assistant, the chatbot will give greetings to students and then provide student statistics from the previous meeting, students are at the level (%) compared to

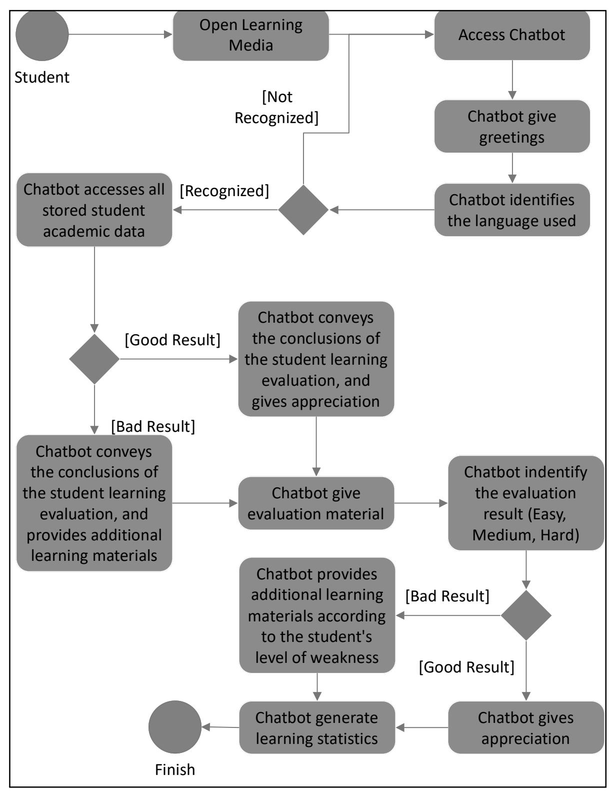

Fig. 5: Chatbot activity diagram

other students in the same class, then the chatbot will present information on the current section of the meeting, after all the material buttons are accessed, the chatbot will "pop up" "Back to evaluate the results of student activities in 1 meeting, if participants cannot solve questions at the easy level, the chatbot will provide additional references from the internet related to easy level questions, as well as medium and hard level questions. After completing all evaluation questions, the chatbot will display a statistical report on how long it took students to read/access the material when compared to other students, students were also presented with information about how long it took the student to complete the evaluation question when compared to other students in the class, this chatbot process can be seen more clearly in Figure 5.

D. Testing Stage Result

The researcher applies the White box and Black box testing methods to test whether the learning media that have been made have worked according to the user's needs. For White box testing, the researcher uses the BLEU score evaluation algorithm, this BLEU metric can be calculated either during model training or after training by counting from the stored one's scales candidate and the reference text is loaded in a different way in each type of calculation. In the case of counting metrics after training the chatbot, the last ones are loaded first trained weights from the set. Then the test data from the two files are loaded. By these, the files are tested.enc and test.dec. The reason for using two files is the need to load the data for the neural network encoder and decoder. After loading all this data, the calculation of the BLEU metric itself begins. The data is stored in the test.dec file is used for the reference text. A candidate, in this case, the chatbot's predicted response is without special tokens. For the second case, which is testing during model training, the test data is not loaded from these two files. The data that is obtained using the train_test_split() function is used. This function selects the data for testing from the data intended for training and these are further used as a reference text. Here, too, the candidate is the same as for calculating the metric from stored weights. Other parts of the BLEU metric calculation are common to both cases, we must calculate the weights to calculate this metric. If the size of the length of the reference text and at the same time the candidate length size is greater than 4 tokens, so the weights are set as follows: weights = (0.25, 0.25, 0.25, 0.25). The results of the White box testing can be seen in the graph below.

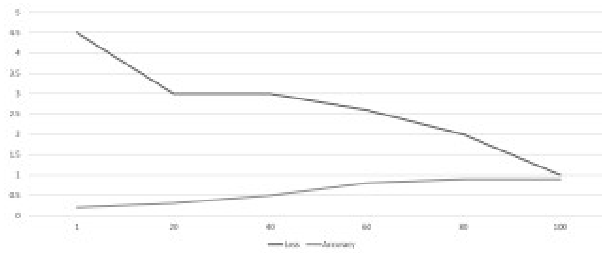

Fig. 6 : The progress of the loss function and accuracy function calculation during the training of the chatbot model

A total of 171,244 data samples were used for training this chatbot model. Of these, 1,000 samples were taken with the reason being to allow use for model validation using the BLEU metric. All models were also trained at 100 epochs and used a lossy approach for training functions and optimization tools.

After doing White box testing and all chatbot functions running as expected, the researcher conducted Black box testing to find out whether the results of each feature still have program errors or bugs. When conducting Blackbox testing, researchers invited 2 media experts and 2 learning content experts. The results of the tests that have been carried out by researchers are presented in Table 1 below.

Table 1 : Black box testing result

No	Scenario	Expected Result	Result
1	User login	Arrived at dashboard page	Valid
2	Learning Materials menu	Meetings 1-14 displayed,	Valid
3	Meeting 1-14 menu	Learning material, animation, and video displayed	Valid
4	Back button	Previous page from learning material displayed	Valid
5	Forward button	Next page from learning material displayed	Valid
6	Play button	Learning video running	Valid
7	Learning evaluation menu	Question for the current meeting displayed	Valid
8	Chatbot button	Enter chatbot page, chatbot give greetings	Valid
9	User interact with chatbot	Chatbot responds and understands questions from users	Valid
10	User ask chatbot for their learning statistics	Chatbot provide learning evaluation and can provide additional learning material	Valid
11	User ask chatbot for evaluation result	Chatbot can identify evaluation result and provides additional learning materials according to the student's level of weakness	Valid
12	Usage guide button	Usage guide page displayed	Valid
13	Edit profile	User can edit and save new profile	Valid
14	Logout button	User logged out	Valid
15	Close button	Program closed	Valid

E. Distribution Stage Result

After conducting Black box and White box testing, the researcher invited 94 students and 1 lecturer to use the learning media that had been made in the teaching and learning process as many as 14 meetings to fill out the questionnaire form and obtain the results which can be seen in Table 2 below.

Table 2 : User acceptance result

Criteria	Neutral (%)	Agree (%)	Stongly Agree (%)
Learning media is useful for thoroughly explaining concepts and case studies in courses.	2 (2)	16 (17)	77 (81)
Learning media has greatly aided my comprehension of the learning material.	7 (7)	12 (13)	76 (80)
This learning media makes it easier for me to find the most recent learning resources	1 (1)	29 (31)	65 (68)
The interface for learning media appears appealing.	0 (0)	32 (34)	63 (66)
My experience when using this learning media is excellent	5 (5)	11 (12)	79 (83)
This learning media suits my needs	2 (2)	15 (16)	78 (82)
The chatbot is very communicative and can provide the necessary information.	0 (0)	2 (2)	93 (88)
This learning media is very easy to use	5 (5)	21 (22)	69 (73)
Mean Value	2.75%	18.3%	77.63%

Based on the data in the table above, the researcher concludes that this chatbot learning media is extremely beneficial to students' comprehension of the learning material. This learning media provides virtual assistants who assist students by providing additional material in areas where students do not understand. Based on the table above, we can also conclude that this learning media has an appealing, simple, and easy-to-use interface. With an average value of 77.63% stated that they strongly agree that this chatbot-assisted learning media is very useful for use during learning, 18,3% agree and 2.75% neutral response.

The effectiveness of chatbot-assisted learning media in basic electronics subjects was analyzed using the T test. The first step in testing the effectiveness of learning media is to conduct a test item test that will be tested on students. The test items carried out are:

A. Question validity test, the instrument in this study uses a multiple choice test, and this validity can be calculated using the product moment correlation coefficient. To find out whether a question is valid or not can be seen by checking the significance value. If the significance value is <0.05 then the question or test item is declared valid. In addition, to determine the validity of the question, you can also see from the comparison between R Count and R Table. If R Count> R Table then the question or test item is declared valid. Based on the results of testing the validity of questions or test items that have been carried out, the results show that there are 24 questions that are declared valid and 6 questions are declared invalid, from testing the validity of the questions there are 22 valid questions that can be used.

B. Reliability test, an instrument that produces the same data when measuring the same object several times. If the instrument has fixed measurement results, it can be said to have high reliability. Reliability represents the understanding that an instrument is reliable enough to be used as a data collection tool because the instrument is good. Reliability indicates an understanding that an instrument can be trusted enough to be used as a data collection tool because the instrument is good. From the calculation of the reliability test with Kuder-Richardson (KR-20), the reliability value is 0.84224. Based on the interpretation of reliability for the coefficient range between 0.700 - 0.89, it is declared to have a high level of reliability.

C. Difficulty test, the level of difficulty can be seen as the ability or ability of students to answer the test given. The level of difficulty of a question item is expressed by a number called the difficulty index. The difficulty index ranges from a value of 0.00 to 1.00. A question with a difficulty index of 0.00 means that the item is too difficult, on the other hand, a question difficulty index close to 1.00 means that the question is too easy. Based on the results of the test of the difficulty of the questions or test items that have been carried out, the data obtained are as in the table above. There are 4 questions with easy category, 14 questions in the medium category, and 4 questions in the difficult category.

D. Differentiated Test Question, to measure the level of ability to distinguish high ability students from low ability students. The discriminating power of a question item is to state how far the ability of the question item is able to distinguish between the test target that knows the answer correctly and the test target that cannot answer the question (test target that answers incorrectly). If the discrimination index of the question is getting closer to the value of 1.00, this means that the discriminating power of the question will be better, and vice versa, if the discrimination index of a question is close to the value of 0.00, the discriminating power of the question is very bad. Based on the results of the test of the differentiation of questions or test items that have been carried out, There are 20 questions with good differentiating power categories and 2 questions with very good differentiating power categories.

E. Normality Test, before conducting the T test, you must first test the normality and homogeneity of the pretest and posttest data. The normality test is used to see if the learning outcomes of the experimental and control classes are normally distributed. The normality test is useful in helping researchers determine which data analysis technique to use. Normally distributed data will be analyzed using parametric statistical techniques. The normality test in this study was calculated using the Kolmogorov-Smirnov formula.

Table 3 : Normality test

	Class	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Learning Outcome	PreTest Experiment Class	.149	18	.200*	.956	18	.518
	PostTest Experiment Class	.196	18	.065	.890	18	.039
	PreTest Control Class	.149	18	.200*	.907	18	.075
	PostTest Control Class	.152	18	.200*	.940	18	.294

*. This is a lower bound of the true significance

a. Lilliefors Significance Correction

From the normality test data in the table above, the significance value is 0.2 for the experimental group pretest, 0.065 for the experimental group posttest, 0.2 for the control group pretest, and 0.2 for the control group posttest. These results show that the significance value obtained is> 0.05, which means that the data is normally distributed.

F. Homogeneity Test, used to decide whether two populations with unknown distributions have the same distribution as each other or are homogeneous. This test is used to determine whether the pretest and posttest data from the experimental group and the control group are homogeneously distributed or not. The homogeneity test can be done using the Levene test. The pairs of hypotheses to be tested are as follows; H0, sample groups come from populations that have the same or homogeneous variance. H1, sample groups come from populations that have different variances or are not homogeneous. The test criteria used in this Levene test are if the Wcount value ≤ 0.05 then the data group is said to have an inhomogeneous variance (H0 rejected). Conversely, if the Wcount value> 0.05, the data group is said to have a homogeneous variance (H0 is accepted).

Table 4 : Homogenity test of variance

		Levene Statistic	df1	df2	Sig.
Learning Outcome	Based on Mean	.113	1	34	.738
	Based on Median	.180	1	34	.673
	Based on Me dian and with adjusted df	.180	1	33.931	.673
	Based on trim med mean	.104	1	34	.748

From the homogeneity test data presented in the table above, it is known that the significance value based on mean is 0.738 where this data> 0.05 then H0 is accepted and H1 is rejected which means that the pretest and posttest data from the experimental group and control group are homogeneously distributed.

G. T-test or independent sample T-test was conducted to determine whether there was a significant difference between the posttest results of the experimental group and the control group. The experimental group as a group that uses the developed learning media while the control group is a group that does not use the chatbot learning media. The hypothesis in this test is H0: there is no difference in student learning outcomes after using chatbot-based learning media in basic electronics subjects. H1: There is a difference in student learning outcomes after using chatbot-based learning media in basic electronics subjects. The basis for decision making in the independent t test is if the Significance value or Sig. (2-tailed) > 0.05, then H0 is accepted and H1 is rejected. If the Significance value or Sig. (2-tailed) <0.05, then H0 is rejected and H1 is accepted.

Table 5 : T Test Results on Student Learning Outcomes

		Levene's Test for Equality		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Learning Outcome	Equal variance assumed	.113	.738	11.000	34	.000	25.000	2.273	20.381	29.619
	Equal variance not assumed	.196	18	11.000	33.864	.000	25.000	2.273	20.381	29.619

Based on the data from the T test results shown in the table above, the sig value is obtained. (2-tailed) of 0.000 <0.05 then H0 is rejected and H1 is accepted so it can be concluded that there is a significant average difference in student learning outcomes between the experimental group and the control group.

5. Conclusion and Future Research Suggestion

The chatbot model that has been applied to this learning media when tested with the BLEU metric gets a result of 0.1117, this shows the chatbot produces good answers and provides credible learning references, the chatbot responds clearly and without many grammatical errors. This research characterizes the neural network and provides general information about the functioning of the chatbot. The chatbot is taught to respond so that it can communicate with the user. However, the purpose of this research is not only to identify the feedback given by the chatbot, this research also implements the chatbot in the LMS and tests its feasibility and user acceptance level. There is a difference in student learning outcomes after using chatbot-based learning media in basic electronics subjects, the utilization of chatbot-based learning media is able to improve student learning outcomes, especially in the student weaknesses section, the chatbot can provide detailed explanations and guide students in solving linear equation problems.

There has been little research on psychological needs in AI for education [38-40]. This study adds to the body of knowledge about AI educational tools and provides proof of their viability. The learning settings studied in AI research are still predominantly non self-regulated learning (SRL) [41-44]. We adopt a different approach, highlighting how AI technology can assist students in becoming more self-sufficient learners [45-50]. It is hoped that the results of this research can form the basis for the development of more effective interactive learning media in the future.

This research is still at the development stage, so further research is needed to see the effect of this media with AI chatbot by implementing it to the learning process directly. Future research should include multi-user testing, a larger number of test subjects [44], and a test to see the effect of academic achievement and student learning motivation when using this LMS.

Acknowledgment

The researchers would like to extend their deepest gratitude to the Direktorat Riset dan Pengabdian Kepada Masyarakat (DRPM) for supporting this study.

References

[1] Batubara, Beby Masitho. "The Problems of the World of Education in the Middle of the Covid-19 Pandemic." Budapest International Research and Critics Institute (BIRCI-Journal): Humanities and Social Sciences 4.1 (2021): 450-457.

[2] Zheng, Changdong, et al. "From active learning to deep reinforcement learning: Intelligent active flow control in suppressing vortex-induced vibration." Physics of Fluids 33.6 (2021): 063607.

[3] Song, Pu, and Xiang Wang. "A bibliometric analysis of worldwide educational artificial intelligence research development in recent twenty years." Asia Pacific Education Review 21 (2020): 473-486.

[4] Zawacki-Richter, Olaf, et al. "Systematic review of research on artificial intelligence applications in higher education–where are the educators?." International Journal of Educational Technology in Higher Education 16.1 (2019): 1-27.

[5] De Vreede, Triparna, Mukhunth Raghavan, and Gert-Jan De Vreede. "Design foundations for AI assisted decision making: A self determination

theory approach." (2021).

[6] Yang, Xi, and Marco Aurisicchio. "Designing conversational agents: A self-determination theory approach." Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. 2021.

[7] Chame, Hendry Ferreira, Fernanda Pinto Mota, and Silvia Silva da Costa Botelho. "A dynamic computational model of motivation based on self-determination theory and CANN." Information Sciences 476 (2019): 319-336.

[8] Alfadda, Hind Abdulaziz, and Hassan Saleh Mahdi. "Measuring students' use of zoom application in language course based on the technology acceptance model (TAM)." Journal of Psycholinguistic Research 50.4 (2021): 883-900.

[9] Cheng, Gary. "Exploring factors influencing the acceptance of visual programming environment among boys and girls in primary schools." Computers in Human Behavior 92 (2019): 361-372.

[10] Chow, Meyrick, et al. "Factors affecting nursing students' intention to use a 3D game to learn field triage skills: a structural equation modelling analysis." BMC nursing 21.1 (2022): 1-9.

[11] Dorfsman, Marcelo, and Gabriel Horenczyk. "The coping of academic staff with an extreme situation: The transition from conventional teaching to online teaching." Education and Information Technologies (2022): 1-23.

[12] Khlaisang, Jintavee, Timothy Teo, and Fang Huang. "Acceptance of a flipped smart application for learning: a study among Thai university students." Interactive Learning Environments 29.5 (2021): 772-789.

[13] Følstad, Asbjørn, and Marita Skjuve. "Chatbots for customer service: user experience and motivation." Proceedings of the 1st international conference on conversational user interfaces. 2019.

[14] Pérez, José Quiroga, Thanasis Daradoumis, and Joan Manuel Marquès Puig. "Rediscovering the use of chatbots in education: A systematic literature review." Computer Applications in Engineering Education 28.6 (2020): 1549-1565.

[15] Coniam, David. "The linguistic accuracy of chatbots: usability from an ESL perspective." Text & Talk 34.5 (2014): 545-567.

[16] Smutny, Pavel, and Petra Schreiberova. "Chatbots for learning: A review of educational chatbots for the Facebook Messenger." Computers & Education 151 (2020): 103862.

[17] Wollny, Sebastian, et al. "Are we there yet?-A systematic literature review on chatbots in education." Frontiers in artificial intelligence 4 (2021): 654924.

[18] Hwang, Gwo-Jen, and Ching-Yi Chang. "A review of opportunities and challenges of chatbots in education." Interactive Learning Environments (2021): 1-14.

[19] Sandu, Nitirajsingh, and Ergun Gide. "Adoption of AI-Chatbots to enhance student learning experience in higher education in India." 2019 18th International Conference on Information Technology Based Higher Education and Training (ITHET). IEEE, 2019.

[20] O'Regan, Gerard. "Joseph Weizenbaum." Giants of Computing. Springer, London, 2013. 263-266.

[21] Singh, Siddhant, and Hardeo K. Thakur. "Survey of various AI Chatbots based on technology used." 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO). IEEE, 2020.

[22] Haristiani, Nuria. "Artificial Intelligence (AI) chatbot as language learning medium: An inquiry." Journal of Physics: Conference Series. Vol. 1387. No. 1. IOP Publishing, 2019.

[23] Agarwal, Ritu, and Mani Wadhwa. "Review of state-of-the-art design techniques for chatbots." SN Computer Science 1.5 (2020): 1-12.

[24] Ahmad, Nahdatul Akma, et al. "Review of chatbots design techniques." International Journal of Computer Applications 181.8 (2018): 7-10.

[25] Oh Song, Hyun, et al. "Deep metric learning via lifted structured feature embedding." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[26] Mondal, Anupam, et al. "Chatbot: An automated conversation system for the educational domain." 2018 International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP). IEEE, 2018.

[27] Noorbehbahani, Fakhroddin, and Ahmad A. Kardan. "The automatic assessment of free text answers using a modified BLEU algorithm." Computers & Education 56.2 (2011): 337-345.

[28] Mohammad Nezhad, Ali, and Hashem Mahlooji. "An artificial neural network meta-model for constrained simulation optimization." Journal of the Operational Research Society 65.8 (2014): 1232-1244.

[29] Bilbao, Imanol, and Javier Bilbao. "Overfitting problem and the over-training in the era of data: Particularly for Artificial Neural Networks." 2017 eighth international conference on intelligent computing and information systems (ICICIS). IEEE, 2017.

[30] Sahronih, Siti, Agung Purwanto, and M. Syarif Sumantri. "The effect of interactive learning media on students' science learning outcomes." Proceedings of the 2019 7th International Conference on Information and Education Technology. 2019.

[31] Sandu, Nitirajsingh, and Ergun Gide. "Adoption of AI-Chatbots to enhance student learning experience in higher education in India." 2019 18th International Conference on Information Technology Based Higher Education and Training (ITHET). IEEE, 2019.

[32] Hamid, Mustofa Abi, Lely Yuliawati, and Didik Aribowo. "Feasibility of Electromechanical Basic Work E-Module as a New Learning Media for Vocational Students." Journal of Education and Learning (EduLearn) 14.2 (2020): 199-211.

[33] Septian, Deris, Yenni Fatman, and Siti Nur. "IMPLEMENTASI MDLC (MULTIMEDIA DEVELOPMENT LIFE CYCLE) DALAM PEMBUATAN MULTIMEDIA PEMBELAJARAN KITAB SAFINAH SUNDA." Jurnal Computech & Bisnis 15.1 (2021): 15-23.

[34] Rauber, Jonas, Matthias Bethge, and Wieland Brendel. "EagerPy: Writing code that works natively with PyTorch, TensorFlow, JAX, and NumPy." arXiv preprint arXiv:2008.04175 (2020).

[35] Drude, Lukas, et al. "NARA-WPE: A Python package for weighted prediction error dereverberation in Numpy and Tensorflow for online and offline processing." Speech Communication; 13th ITG-Symposium. VDE, 2018.

[36] Syaikhuddin, Muhammad Miftakhul, et al. "Conventional software testing using white box method." Kinetik: game technology, information system, computer network, computing, electronics, and control (2018): 65-72.

[37] Sefrika, S., and Enok Tuti Alawiah. "Acceptance of Video Conference Technology as a Distance Learning Media with the TAM Method." IJISTECH (International Journal of Information System and Technology) 4.1 (2020): 465-470.

[38] Al Shamsi, Jawaher Hamad, Mostafa Al-Emran, and Khaled Shaalan. "Understanding key drivers affecting students' use of artificial intelligence-based voice assistants." Education and Information Technologies 27.6 (2022): 8071-8091.

[39] Gado, Sabrina, et al. "Artificial intelligence in psychology: How can we enable psychology students to accept and use artificial intelligence?." Psychology Learning & Teaching 21.1 (2022): 37-56.

[40] Seufert, Sabine, Josef Guggemos, and Michael Sailer. "Technology-related knowledge, skills, and attitudes of pre-and in-service teachers: The current situation and emerging trends." Computers in Human Behavior 115 (2021): 106552.

[41] Bedenlier, Svenja, et al. "Facilitating student engagement through educational technology in higher education: A systematic review in the

field of arts and humanities." Australasian Journal of Educational Technology 36.4 (2020): 126-150.

[42] Jansen, Renée S., et al. "Self-regulated learning partially mediates the effect of self-regulated learning interventions on achievement in higher education: A meta-analysis." Educational Research Review 28 (2019): 100292.

[43] Carter Jr, Richard Allen, et al. "Self-regulated learning in online learning environments: strategies for remote learning." Information and Learning Sciences 121.5/6 (2020): 321-329.

[44] Azevedo, Roger, and Dragan Gašević. "Analyzing multimodal multichannel data about self-regulated learning with advanced learning technologies: Issues and challenges." Computers in Human Behavior 96 (2019): 207-210.

[45] Sriadhi, S., et al. "Effectiveness of Augmented Reality-Based Learning Media for Engineering-Physics Teaching." International Journal of Emerging Technologies in Learning 17.5 (2022).

[46] Mendoza, Muhammad Dominique, et al. "PENGARUH PENGGUNAAN MEDIA SOSIAL DALAM PENDIDIKAN TERHADAP PRESTASI AKADEMIK MAHASISWA." Jurnal Teknologi Pendidikan (JTP) 15.2 (2022): 68-80.

[47] Gultom, Syamsul, et al. "Exploring the perceptions of student-athletes leading towards physical activity and perceived performance." Revista de Psicología del Deporte (Journal of Sport Psychology) 30.4 (2021): 201-215.

[48] Mendoza, M. D., O. Y. Hutajulu, and R. Salman. "Community constraints analysis in the use of solar power plants in Indonesia." Journal of Physics: Conference Series. Vol. 2193. No. 1. IOP Publishing, 2022.

[49] Sriadhi, S., et al. "Virtual-laboratory based learning to improve students' basic engineering competencies based on their spatial abilities." Computer Applications in Engineering Education 30.6 (2022): 1857-1871.

[50] Sriadhi, S., et al. “Web-Based Virtual Laboratory Development for Basic Practicums in Science and Technology.” TEM Journal, 2022, pp. 396–402, https://doi.org/10.18421/tem111-50.