Jurnal Teknologi Informasi dan Pendidikan
Volume 17, No. 2, September 2024
https://doi.org/10.24036/jtip.v17i2.853

P.ISSN: 2086 – 4981
E.ISSN: 2620 – 6390
tip.ppj.unp.ac.id

Effectiveness of AI-Based Chatbots as Virtual Tutors in Basic Programming Education

Arif Setiawan[1*]✉, Muhammad Fikri Irfanuddin[1], Ryan Rizki Adhisa[1]
[1]Informatics Engineering Education, Universitas Muhammadiyah Surakarta, Surakarta, Indonesia
Corresponding Author: arif.setiawan@ums.ac.id

Article Information

Article history:

No. 853
Rec. May 15, 2024
Rev. December 14, 2024
Acc. December 16, 2024
Pub. December 23, 2024
Page. 482 – 492

Keywords:

- *Ai-based chatbot*
- *Basic programming education*
- *Fine-tunning chatbot*
- *Open AI*

ABSTRACT

The integration of AI-based chatbots as virtual tutors in basic programming education addresses the challenges posed by the diverse backgrounds of students and the complexity of the course material. This integration aims to provide personalized assistance and explanations tailored to individual student needs and understanding levels. The research methodology employed in this study adopts a Research and Design (RnD) approach, utilizing the 4D development model, which encompasses the stages of Definition, Design, Development, and Deployment. Through this methodology, the study aims to develop and evaluate the effectiveness of the chatbot in enhancing students' learning experiences. The Lighthouse testing revealed performance scores of 98 for performance, 71 for accessibility, 96 for best practices, and 80 for SEO. These results indicate that the application adheres to high technical standards, ensuring its quality and usability. The findings further indicate a significant increase in students' self-learning activities, as evidenced by a score of 79.4% in the "Very Good" category in the Likert scale evaluation. This suggests that the integration of AI technology in basic programming education holds great potential for enhancing accessibility and effectiveness in learning, while also enriching students' learning experiences by providing adaptive and personalized learning resources. Overall, the findings underscore the importance of leveraging AI technology to address the challenges faced in programming education, thereby fostering a more inclusive and effective learning environment.

How to Cite:

Setiawan, A., et al. (2024). Effectiveness of AI-Based Chatbots as Virtual Tutors in Basic Programming Education. Jurnal Teknologi Informasi Dan Pendidikan, 17(2), 482-492. https://doi.org/10.24036/jtip.v17i2.853

1. INTRODUCTION

The integration of AI-based chatbots as virtual tutors in education is a rapidly growing phenomenon [1]. In the era of digital transformation and the Fourth Industrial Revolution, the need for differentiated and personalized learning methods is becoming increasingly urgent [2]. AI-based chatbots offer various benefits, including more personalized, adaptive learning, and student learning data analysis [3]. This technology enables more dynamic learning methods, allowing materials to be tailored to the needs and interests of individual students.

Basic programming is a mandatory course for first-year students in Informatics study programs. This course is crucial as it forms the foundation for more complex subjects. However, students face challenges due to their diverse backgrounds and varying levels of computer proficiency. Some are already familiar with computers, while others have never operated one [4]. Additionally, basic programming is inherently complex, requiring an understanding of abstraction, logic, and algorithms, further complicating it for students without a programming background [5]. A trial test for first-year computer science students showed an average score of only 22.89 out of 110 points [6], indicating the difficulty of the course.

Research on integrating virtual tutors in education has increased in recent years. However, there is a lack of studies specifically examining the impact of virtual tutors on student learning outcomes. Most research focuses on technological aspects and implementation [7] [8] [9], neglecting how this technology affects student achievement. Conventional teaching methods have proven effective in some cases, but they often fail to meet the diverse needs of students [10]. Therefore, further research is needed to focus not only on technology implementation but also on how virtual tutors impact learning outcomes to achieve optimal results.

Artificial Intelligence (AI) has been widely utilized in education, including programming education from basic levels to higher education. One application of AI is the use of machine learning to aid in teaching and evaluation in K-12 education [11]. In higher education, AI has been extensively used in programming learning processes [12]. AI-based multimedia technology has also been developed as an alternative teaching method to support more effective learning [13].

The advancement of generative AI technology shows significant progress. Generative models such as GANs, VAEs, and Transformers have been used in various applications, including image synthesis and text generation [14]. In education, AI-based chatbots have become effective tools as virtual tutors. For example, ChatGPT has been employed in various subjects, including programming [15]. AI chatbots provide informative feedback to students and can adapt to different learning needs. For instance, BookBuddy is a chatbot used to turn reading materials into interactive English lessons [16]. Additionally,

chatbots like e-Java are used in programming education to provide feedback to students solving Java programming problems [17].

AI-based chatbots hold significant potential, especially in education. However, there are challenges to address, such as the risk of misinformation. For example, research on the use of ChatGPT in dermatology found that, while the chatbot could support health literacy, it also had the potential to mislead patients and provide incorrect diagnoses [18]. On the other hand, AI chatbots offer flexibility in knowledge customization. Fine-tuning features allow AI models to be adapted to user needs. For instance, the GLIDE model, originally designed for text-based image generation, demonstrated accurate results in cancer and oncology image generation after fine-tuning [19]. Similarly, GPT-3's fine-tuning enhanced the output quality in generating design concepts in biology [20].

The development of AI-based chatbots in programming education shows great promise. For example, Python-Bot is designed to help novice programmers understand the basic syntax and semantics of the Python programming language [21]. Furthermore, research has integrated AI chatbots into introductory programming courses, enabling students to receive real-time material and feedback from instructors [22]. Unlike previous studies, the AI chatbot in this research will be developed using fine-tuning based on the learning module documents taught in the Informatics Engineering Education program at Universitas Muhammadiyah Surakarta. The novelty of this research lies in its focus on examining the impact of AI chatbots on student motivation. Most previous studies have focused on chatbot development technology rather than its impact on student motivation. Therefore, this research aims to provide new insights into how fine-tuning AI chatbots can be tailored to course needs and how it impacts student learning activities.

2. RESEARCH METHOD

This research falls under the category of Research and Design (RnD). RnD is a research approach used to develop new products or systems [23]. In the context of education, RnD is employed to develop learning materials, teaching methods, or educational technologies. RnD has been utilized in various educational research projects, such as the development of Android-based games for environmental education [24] and the development of mathematics learning media with an ethnomathematics approach [25]. The stages of this research will follow the 4D development model, which consists of the stages of Define, Design, Development, and Dissemination.

2.1. Define

The Define stage focus on identifying and analyzing needs. The first step involves identifying the problems faced by students, such as difficulties in understanding concepts or the need for personalized learning. Subsequently, a needs analysis is conducted through

P.ISSN: 2086 – 4981
E.ISSN: 2620 – 6390
tip.ppj.unp.ac.id

surveys or interviews with stakeholders, including lecturers and experts in the field of artificial intelligence. A literature review is also performed to understand existing solutions and identify gaps that this research can address.

2.2. Design

The Design stage will focus on developing an AI-based chatbot to be integrated into the learning process. This design will leverage the fine-tuning feature provided by OpenAI to ensure that the chatbot aligns with the educational needs and context of Universitas Muhammadiyah Surakarta [26]. The knowledge base will utilize practical module documents and learning modules in PDF format, totaling over 200 pages. The fine-tuning process flow is illustrated in Figure 1.

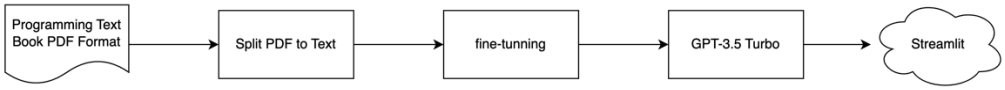

Figure 1. Fine-tunning Process Flow

Additionally, the designed chatbot will be accessible via a web browser, making it convenient for students to access anytime and anywhere without the need to download additional applications. The wireframe design of the chatbot homepage is illustrated in Figure 2.

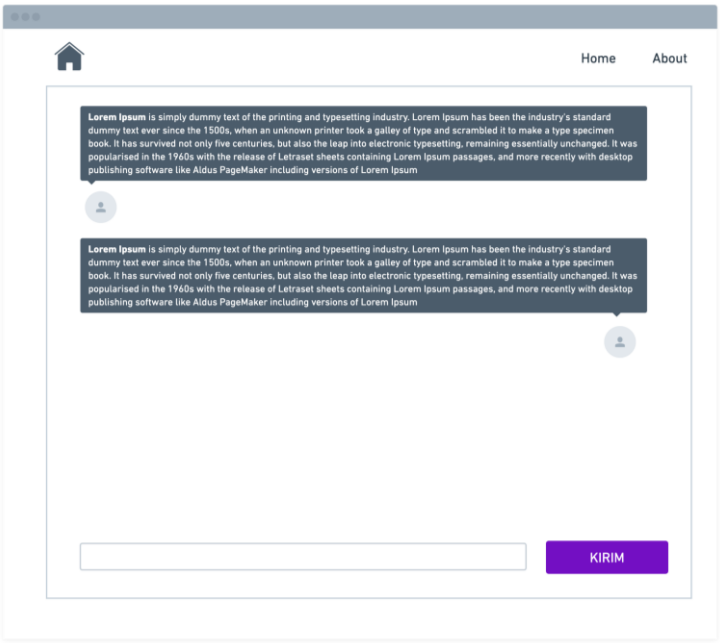

Figure 2. Chatbot UI Design

2.3. Development

The Development Stage involves the production and testing of the designed AI-based chatbot. The production process will utilize the Python programming language and the Streamlit platform, known for its flexibility and efficiency in web application development. Once the chatbot is successfully developed, the next phase involves a series of tests to ensure its effectiveness and efficiency in an educational context.

The testing will be conducted in three stages. First, the performance of the chatbot application will be tested using the Lighthouse tool. Second, expert evaluation will involve three lecturers from the Informatics Engineering Education department. These experts will assess the chatbot in terms of information accuracy, responsiveness, and relevance to the introductory programming course. Third, student testing will involve 25 students from the Algorithm and Programming course, Informatics Engineering Education program at Universitas Muhammadiyah Surakarta. This phase will span four sessions over four weeks.

Data collection will be conducted using questionnaires. These questionnaires will serve as instruments to evaluate the learning activities and experiences of students after interacting with the chatbot. The questionnaire, adapted from the study by Avci and Ersoy, is specifically designed to assess learning motivation in programming courses [27]. It employs a Likert scale and consists of 15 items measuring six aspects: attitude, goal, clarity of direction, reward, punishment, social pressure, and competition.

The evaluation instrument employed a Likert scale with values ranging from 1 to 5, representing respondents' levels of agreement with the questionnaire statements. These values correspond to: (1) strongly disagree, (2) disagree, (3) neutral, (4) agree, and (5) strongly agree. This scale was applied to assess three key aspects: learning display, learning content, and application use.

2.4. Dissemination

The final stage, Dissemination, focuses on the implementation and distribution of the research findings. During the implementation phase, monitoring and evaluation will be conducted to optimize and improve the chatbot features.

3. RESULTS AND DISCUSSION

3.1. Result

The chatbot application has been successfully developed using Python and Streamlit platforms. This combination allows for efficient and user-friendly application development. The chatbot has significant capabilities in understanding the context provided by users. Utilizing OpenAI's API, the chatbot can recognize conversational patterns, analyze the

P.ISSN: 2086 – 4981
E.ISSN: 2620 – 6390
tip.ppj.unp.ac.id

meaning of user questions or statements, and provide relevant and accurate responses. This enhances the naturalness and effectiveness of interactions with the chatbot.

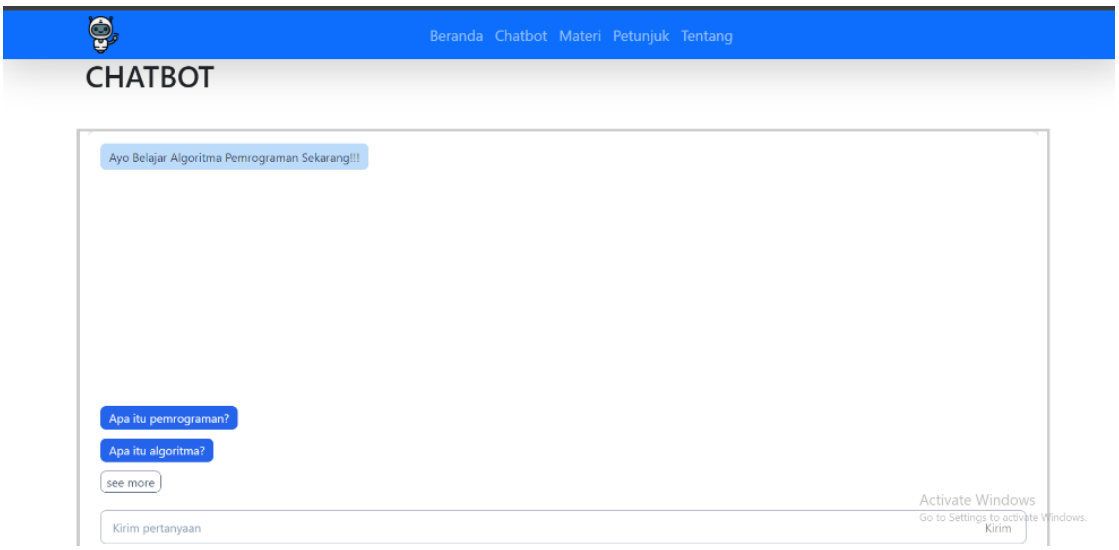

Figure 3. Chatbot User Interface

Figure 3 illustrates the simple and intuitive interface of the chatbot application. Users can type their questions or requests into the provided input field, and the chatbot will respond based on its knowledge base. The user-friendly interface facilitates interaction with the chatbot without requiring extensive technical knowledge. The chatbot employs natural language processing technology to analyze user queries and provide relevant and accurate answers.

Figure 4. Chatbot response when the message matches the lecture material

Volume 17, No. 2, September 2024
https://doi.org/10.24036/jtip.v17i2.853

The use of generative AI enables the chatbot to more effectively understand the context of user questions. This allows the chatbot to provide more relevant and in-depth responses, aligned with the material taught by instructors. Figure 4 demonstrates how the chatbot delivers appropriate responses when user questions correspond to the course content. With generative AI technology, the chatbot can analyze questions more comprehensively, identify implicit contexts, and offer detailed answers.

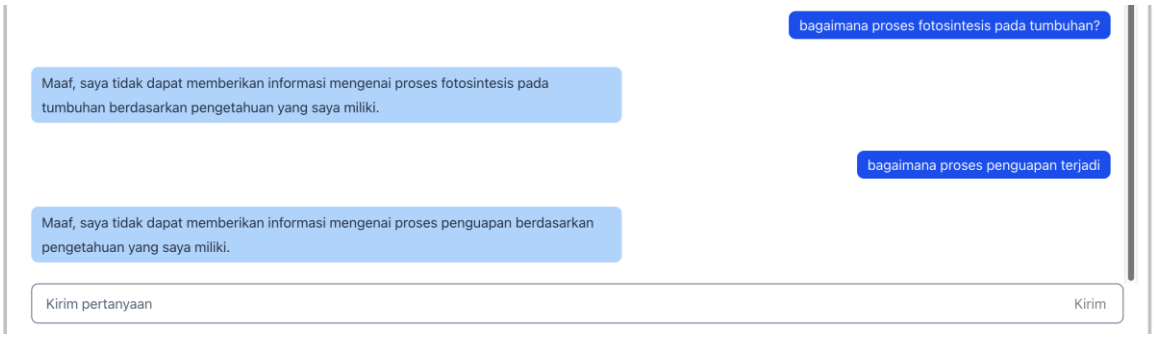

Figure 5. Chatbot response when the message does not matches the lecture material

While the chatbot can understand the context of user questions, its responses rely on the pre-established knowledge base. Figure 5 illustrates an example of the chatbot's response when a user query is unrelated to programming. This highlights the importance of having a comprehensive and relevant knowledge base to ensure accurate responses from the chatbot. Although the chatbot's ability to grasp context is crucial, the accuracy of its responses remains dependent on the depth and completeness of the utilized knowledge base.

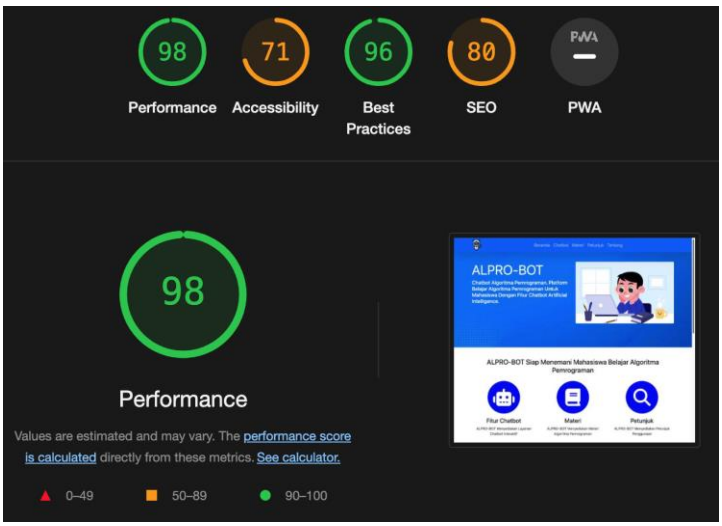

Figure 6. Lighthouse test result

P.ISSN: 2086 – 4981
E.ISSN: 2620 – 6390
tip.ppj.unp.ac.id

The first test was conducted using the Lighthouse tool from the Chrome Browser application to measure the application's performance. The test results can be seen in Figure 6, showing performance scores of 98, accessibility 71, best practices 96, and SEO 80. These scores provide a comprehensive overview of the overall application quality, including how well the application can be accessed, implements best practices, and is optimized for search engines.

The next testing step involved three lecturers from the Informatics Engineering Department, who evaluated the application using a Likert questionnaire. This questionnaire covered three aspects of questions: learning display, learning content, and application usage. The test results showed an overall score of 87%, placing it in the "highly satisfactory" category that presented on table 1. Evaluation from these experts provided valuable insights into how well the application meets expected learning standards, including user satisfaction with the interface, content, and ease of application usage.

Table 1 Expert Testing Result

Aspect	Average Score %	Category
Learning Display	86	Excellent
Learning Content	90	Excellent
Application Use	85	Good

The final trial was conducted with 25 first-year students of the Educational Technology Informatics Study Program. They were introduced to the chatbot website learning media for about 4 weeks. The purpose of this trial was to evaluate its impact on improving students' independent learning activities. In the final week, students were asked to fill out a questionnaire consisting of 15 items to assess the increase in chatbot usage activities. The results of the usage trial showed a score of 79.4%, which falls into the "Very Good" category according to the Likert scale calculation formula. This indicates that the chatbot website successfully significantly improved students' independent learning activities.

3.2. Discussion

The chatbot application has been successfully developed using the Python and Streamlit platforms. This combination enables the efficient development of a user-friendly application. The chatbot exhibits significant capabilities in understanding user context and providing relevant and accurate responses.

Application testing was conducted using various methods. Firstly, performance testing using the Lighthouse tool from the Chrome Browser application provided a comprehensive overview of the overall application quality. The test results indicated excellent scores for performance, accessibility, best practices, and SEO. Subsequently, testing was performed by three lecturers from the Department of Educational Technology

Informatics using a Likert questionnaire, covering aspects of learning display, learning content, and application usage. The expert testing yielded a score of 87%, placing it in the highly satisfactory category. Finally, a trial involving 25 first-year students of the Educational Technology Informatics Study Program assessed the improvement in students' independent learning activities using the chatbot. The questionnaire results obtained a score of 79.4%, classified as very good.

Despite the chatbot's capability to understand context and provide relevant responses, a major challenge lies in the limitations of the pre-established knowledge base. The quality of the chatbot's responses heavily relies on the depth and completeness of this knowledge base. The implication of chatbot usage in education is the enhancement of efficiency and effectiveness in students' self-directed learning [28]. However, attention is needed in the development and maintenance of a relevant knowledge base. Additionally, the adoption of chatbots as learning aids can improve the effectiveness and efficiency of learning, enriching students' learning experiences. However, further research is needed on privacy and data security aspects in their use.

The development of this chatbot application contributes to improving the accessibility of learning and interaction between students and learning materials. The next development direction is to enhance the depth and completeness of the chatbot's knowledge base and integrate new features to enhance user experience and learning outcomes. Furthermore, consideration should also be given to integration with other learning systems such as Learning Management Systems (LMS) to create a holistic learning ecosystem [29].

4. CONCLUSION

This study presents findings regarding the integration of AI-based chatbots in basic programming education for first-year students in the Educational Technology Informatics Study Program. Testing of the chatbot application indicates that the use of AI technology significantly enhances students' self-directed learning activities, as evidenced by the testing results yielding a score of 79.4%, categorized as "Very Good." Moreover, performance testing using the Lighthouse tool revealed high-quality scores of 98 for performance, 71 for accessibility, 96 for best practices, and 80 for SEO, demonstrating the chatbot’s adherence to technical standards. Despite limitations concerning the knowledge base underlying the chatbot’s responses, these findings highlight its substantial contribution to improving learning accessibility and effectiveness. Future development should focus on expanding the knowledge base's depth and completeness and integrating the chatbot with other learning systems such as Learning Management Systems (LMS) to establish a more holistic and integrated learning ecosystem.

P.ISSN: 2086 – 4981
E.ISSN: 2620 – 6390
tip.ppj.unp.ac.id

ACKNOWLEDGEMENTS

Author express gratitude to the Informatics Engineering Education Program at Universitas Muhammadiyah Surakarta for their support. This research was fully supported by the RisetMu grant program from the Majelis Diktilitbang of Muhammadiyah number 0258.066/I.3/D/2024.

REFERENCES

[1] C. K. Y. Chan, "A comprehensive AI policy education framework for university teaching and learning," *Int J Educ Technol High Educ*, vol. 20, no. 1, p. 38, Jul. 2023, doi: 10.1186/s41239-023-00408-3.

[2] F. Yustiasari Liriwati, "Transformasi Kurikulum; Kecerdasan Buatan untuk Membangun Pendidikan yang Relevan di Masa Depan," *J.Ihsan*, vol. 1, no. 2, pp. 62–71, Jul. 2023, doi: 10.61104/ihsan.v1i2.61.

[3] A. D. O. Silva and D. D. S. Janes, "Exploring the Role of Artificial Intelligence in Education: A Comprehensive Perspective," *Rev. Artif. Intell. Educ*, vol. 1, no. 00, p. e05, Dec. 2020, doi: 10.37497/rev.artif.intell.education.v1i00.5.

[4] Josephat O Oroma, H. Wanga, and F. Ngumbuke, "Challenges Of Teaching And Learning Computer Programming In Developing Countries: Lessons From Tumaini University," 2012, doi: 10.13140/2.1.3836.6407.

[5] J. Bennedsen and M. E. Caspersen, "Failure rates in introductory programming," *SIGCSE Bull.*, vol. 39, no. 2, pp. 32–36, Jun. 2007, doi: 10.1145/1272848.1272879.

[6] M. McCracken *et al.*, "A multi-national, multi-institutional study of assessment of programming skills of first-year CS students," *SIGCSE Bull.*, vol. 33, no. 4, pp. 125–180, Dec. 2001, doi: 10.1145/572139.572181.

[7] F. French, D. Levi, C. Maczo, A. Simonaityte, S. Triantafyllidis, and G. Varda, "Creative Use of OpenAI in Education: Case Studies from Game Development," *MTI*, vol. 7, no. 8, p. 81, Aug. 2023, doi: 10.3390/mti7080081.

[8] J. Huang, S. Saleh, and Y. Liu, "A Review on Artificial Intelligence in Education," *Acad. J. Interdiscip. Stud.*, vol. 10, no. 3, p. 206, May 2021, doi: 10.36941/ajis-2021-0077.

[9] L. Chen, P. Chen, and Z. Lin, "Artificial intelligence in education: A review," *Ieee Access*, vol. 8, pp. 75264–75278, 2020.

[10] C. K. Y. Chan and W. Hu, "Students' voices on generative AI: perceptions, benefits, and challenges in higher education," *Int J Educ Technol High Educ*, vol. 20, no. 1, p. 43, Jul. 2023, doi: 10.1186/s41239-023-00411-8.

[11] M. Tedre *et al.*, "Teaching Machine Learning in K–12 Classroom: Pedagogical and Technological Trajectories for Artificial Intelligence Education," *IEEE Access*, vol. 9, pp. 110558–110572, 2021, doi: 10.1109/ACCESS.2021.3097962.

[12] The Korean Society of Culture and Convergence, G. Lee, S. Huh, H. Chee, M. Kim, and M. Kim, "Research on the Development of Teaching and Learning Support Models for the Utilization of AI-based Learning Assistance Systems in Universitie," *Korean Soc Cult Converg*, vol. 45, no. 9, pp. 181–191, Sep. 2023, doi: 10.33645/cnc.2023.09.45.09.181.

[13] Y. Zhang, "Development and Application of Artificial Intelligence Multimedia Technology Based on Big Data," *Mobile Information Systems*, vol. 2022, pp. 1–10, Jan. 2022, doi: 10.1155/2022/2073091.
[14] Tapabrato Bandyopadhyay, Srinjoy Saha, and Debrupa Pal, "Beyond Imitation: Exploring Novelty in Generative AI," *IJARSCT*, pp. 472–475, Sep. 2023, doi: 10.48175/IJARSCT-13070.
[15] E. Chen, R. Huang, H.-S. Chen, Y.-H. Tseng, and L.-Y. Li, "GPTutor: a ChatGPT-powered programming tool for code explanation".
[16] S. Ruan *et al.*, "BookBuddy: Turning Digital Materials Into Interactive Foreign Language Lessons Through a Voice Chatbot," in *Proceedings of the Sixth (2019) ACM Conference on Learning @ Scale*, Chicago IL USA: ACM, Jun. 2019, pp. 1–4. doi: 10.1145/3330430.3333643.
[17] S. H. Mad Daud, "e-JAVA Chatbot for Learning Programming Language: A Post-Pandemic Alternative Virtual Tutor," *IJETER*, vol. 8, no. 7, pp. 3290–3298, Jul. 2020, doi: 10.30534/ijeter/2020/67872020.
[18] E. Porter, M. Murphy, and C. O'Connor, "Chat GPT in dermatology: Progressive or problematic?," *Acad Dermatol Venereol*, vol. 37, no. 7, Jul. 2023, doi: 10.1111/jdv.19174.
[19] J. N. Kather, N. Ghaffari Laleh, S. Foersch, and D. Truhn, "Medical domain knowledge in domain-agnostic generative AI," *npj Digit. Med.*, vol. 5, no. 1, p. 90, Jul. 2022, doi: 10.1038/s41746-022-00634-5.
[20] Q. Zhu, X. Zhang, and J. Luo, "Biologically Inspired Design Concept Generation Using Generative Pre-Trained Transformers," 2022, doi: 10.48550/ARXIV.2212.13196.
[21] "Python-Bot." Accessed: Oct. 22, 2023. [Online]. Available: https://webchat.snatchbot.me/39031b004171a74999c5474f3f41b0e2cb6302679b027223205818971 96c88b1
[22] M. Verleger and J. Pembridge, "A Pilot Study Integrating an AI-driven Chatbot in an Introductory Programming Course," in *2018 IEEE Frontiers in Education Conference (FIE)*, San Jose, CA, USA: IEEE, Oct. 2018, pp. 1–4. doi: 10.1109/FIE.2018.8659282.
[23] Sutama, *Metode Penelitian Pendidikan*. Surakarta: CV. Jasmine, 2019.
[24] H. Rahmayanti, V. Oktaviani, and Y. Syani, "Development of sorting waste game android based for early childhood in environmental education," *J. Phys.: Conf. Ser.*, vol. 1434, no. 1, p. 012029, Jan. 2020, doi: 10.1088/1742-6596/1434/1/012029.
[25] A. Irawan, M. Lestari, and W. Rahayu, "Konsep Etnomatematika Batik Tradisional Jawa Sebagai Pengembangan Media Pembelajaran Matematika," *Scholaria*, vol. 12, no. 1, pp. 39–45, Jan. 2022, doi: 10.24246/j.js.2022.v12.i1.p39-45.
[26] "OpenAI Platform." Accessed: Oct. 25, 2023. [Online]. Available: https://platform.openai.com
[27] A. Ümmühan and H. ERSOY, "The adaptation of learning motivation in computer programming courses scale into Turkish: the study of validity and reliability," *Journal of Higher Education and Science*, vol. 8, no. 1, pp. 073–081, 2018.
[28] X. Deng and Z. Yu, "A Meta-Analysis and Systematic Review of the Effect of Chatbot Technology Use in Sustainable Education," *Sustainability*, 2023, doi: 10.3390/su15042940.
[29] M. Shilowaras and N. Jusoh, "Implementing Artificial Intelligence Chatbot in Moodle Learning Management System," *Engineering, Agriculture, Science and Technology Journal (EAST-J)*, 2022, doi: 10.37698/eastj.v1i1.122.

P.ISSN: 2086 – 4981
E.ISSN: 2620 – 6390
tip.ppj.unp.ac.id