

Application of Deep Learning to Improve Mathematical Literacy in HOTS at MTs Al-Ihsan

Anzil Rahmawati, Khusnul Khotimah

Universitas KH. A. Wahab Hasbullah

Email: khusnulhotimah@unwaha.ac.id

ABSTRACT

This study aims to analyze the implementation of the deep learning approach in improving students' mathematical literacy in solving Higher Order Thinking Skills (HOTS) problems at MTs Al-Ihsan, Jombang. The research employed a quasi-experimental design with a pretest-posttest control group model. The subjects consisted of two classes, VII B as the experimental group and VII C as the control group, each comprising 27 students. The experimental class received mathematics instruction using a deep learning approach that integrated the principles of mindful, meaningful, and joyful learning, while the control class was taught using conventional methods. Data were collected through validated pretest and posttest instruments and analyzed using SPSS. The results of the normality and homogeneity tests showed that the data met the assumptions for parametric analysis (Sig. > 0.05). The independent sample t-test revealed a significant difference between the posttest results of the control and experimental groups (Sig. = 0.000 < 0.05). The N-Gain analysis showed that the experimental class achieved an average score of 0.76 (76.08%), categorized as high, while the control class obtained 0.43 (43.23%), categorized as medium. These findings indicate that the deep learning approach effectively enhances students' mathematical literacy, particularly in solving HOTS-based problems.

Keywords: *Deep Learning; Mathematical Literacy; HOTS; Mathematics Education; Artificial Intelligence.*

INTRODUCTION

With the rapid flow of information in the era of globalization, Indonesia's education system must continuously improve to enhance the quality of human resources capable of competing globally. The UNESCO (2011) Global Education for All (EFA) Monitoring Report noted a decline in the quality of education in Indonesia, with the country's ranking dropping from 65th to 69th out of 127 countries. This condition reflects the ongoing challenges faced by the education system, one of which is the low ability of students to solve problems that require higher-order thinking skills.

Education in Indonesia still faces various challenges in preparing human resources capable of competing in the global era. The UNESCO (2019) report indicates that the quality of education in Indonesia remains relatively low, as students struggle to solve problems that require higher-order thinking skills. Mathematical literacy is one of the 21st-century skills that plays an important role in developing students' critical and creative thinking abilities (Pratiwi & Nugroho, 2020).

Literacy serves as a key competency underpinning 21st-century skills, especially in developing critical and creative thinking abilities. According to Pratiwi and Nugroho (2020), strong literacy skills enable students to access, comprehend, and effectively use information in the learning process. Mathematics, with its complexity and logical structure, is one of the main subjects that foster these abilities through problems based on Higher Order Thinking Skills (HOTS).

Observations at MTs Al-Ihsan show that seventh-grade students still experience difficulties in solving HOTS questions. This finding is consistent with the PISA (2020) report, which states that Indonesian students' ability to solve non-routine problems is still low (Erfan & Ratu, 2020). Therefore, learning innovations such as the deep learning approach are needed to improve students' mathematical literacy. The deep learning approach emphasizes adaptive, meaningful learning supported by technology (Nugroho et al., 2022; Hidayati & Sari, 2023). This study aims to determine the effectiveness of applying

the deep learning approach in improving students' mathematical literacy in solving HOTS problems.

Based on interviews with teachers at MTs Al-Ihsan, eighth-grade students at the school still face difficulties in solving mathematics problems of a Higher Order Thinking Skills (HOTS) nature. Lestari and Firmansyah (2021) revealed that many students struggle to understand basic concepts, apply formulas, and write units correctly, particularly in topics such as the Pythagorean theorem. These errors indicate the need for a learning approach that can address the demands of higher-level thinking.

Findings from the 2020 Program for International Student Assessment (PISA) also showed that Indonesian students continue to perform poorly in solving non-routine problems that require critical reasoning (Erfan & Ratu, 2020). One contributing factor is the continued use of traditional teaching methods that do not fully support the development of analytical and creative thinking. Therefore, it is necessary to implement learning strategies that are innovative and contextual.

The integration of technology in education must be accompanied by the development of a curriculum that is responsive to the changing times. Darmawan (2023) indicated that artificial intelligence technologies, such as deep learning, can create more efficient learning processes tailored to individual learning characteristics. Therefore, it is important to explore how this model can be optimally implemented in school settings.

This study is motivated by the low mathematical literacy of students at MTs Al-Ihsan in solving HOTS problems. This condition poses a significant challenge in strengthening higher-order thinking skills at the secondary education level. The issue needs to be addressed through a learning approach that can optimally facilitate students' cognitive development.

The implementation of deep learning was chosen due to its ability to personalize learning based on an analysis of student performance data. This technology is expected to provide guidance tailored to the level of difficulty faced by students, while also enhancing learning effectiveness. With deep learning, students can receive adaptive learning that meets their individual needs. The conceptual framework of this study posits that students' mathematical literacy can be improved through the application of a deep learning approach integrated into the learning process. Using a quantitative design, the researcher will measure the effectiveness of this approach through pre-test and post-test results on HOTS problems. The study is expected to address the main issues and provide strategic recommendations for educators in integrating technology into mathematics instruction. Thus, the implementation of deep learning is anticipated to impact not only academic outcomes but also the overall development of students' thinking skills.

The researcher states that mathematical literacy involves a complex thinking process, ranging from identifying problems, constructing mathematical models, performing calculations, to drawing conclusions and communicating results. This ability enables students to navigate a modern world filled with numerical information, graphs, and statistical data that require deep understanding to interpret accurately (Siregar & Lestari, 2022). Mathematical literacy demands more than just the ability to answer questions; it requires mastery of context and the meaningful application of mathematics.

Many experts also emphasize that mathematical literacy is at the core of 21st-century competencies because it integrates cognitive, affective, and conative aspects within the mathematical thinking process. This skill allows students to connect mathematical knowledge to real-life situations, such as managing personal finances, interpreting data graphs, or evaluating risks in daily life (Nugroho & Fatimah, 2020). Thus, mathematical literacy is not merely an academic achievement but an essential life skill for surviving and thriving in the complexities of the modern world.

The indicators for assessing students' mathematical literacy based on the PISA (2018) Mathematics Framework are as follows:

Ability to Formulate (Formulate) the term "formulate" in the definition of mathematical literacy refers to a student's ability to design solutions for problems presented in contextual forms. In the process of mathematically formulating a situation, students determine which parts are important for analyzing, organizing, and solving the problem. Students can translate real-world problems into mathematical models and solve them using mathematical structures and representations.

Ability to Employ (Employ) the term "employ" in the definition of mathematical literacy refers to an individual's ability to apply mathematical concepts, facts, procedures, and reasoning to draw mathematical conclusions. In the process of using concepts, facts, procedures, and reasoning to solve problems, students perform the necessary mathematical procedures to obtain results and find solutions.

Ability to Interpret (Interpret) the term "interpret" in the definition of mathematical literacy focuses on a student's ability to consider mathematical solutions, results, or conclusions and interpret them in the

context of real-life problems. This involves translating mathematical solutions or reasoning back into the problem context and determining whether the results are reasonable. Students engaged in this process may also be required to construct and communicate explanations and arguments within the problem context, reflecting on both the modeling process and the outcomes.

This study is motivated by the low mathematical literacy of eighth-grade students at MTs Al-Ihsan, which affects their success in solving HOTS problems. Based on this phenomenon, the author proposes the title: "Application of Deep Learning to Improve Mathematical Literacy in HOTS at MTs Al-Ihsan."

This study originates from the problem of students' low mathematical literacy in solving HOTS problems. This issue underlies the formulation of the research questions as follows:

1. How is the deep learning model implemented to improve mathematical literacy skills in HOTS problems at MTs Al-Ihsan?
2. How effective is the deep learning model in enhancing mathematical literacy skills in HOTS problems at MTs Al-Ihsan?

This study is expected to provide both theoretical and practical benefits for the field of education. Theoretically, it can enrich the literature in mathematics education, particularly regarding the use of deep learning technology in HOTS-based learning. Pratama and Lestari (2022) state that the utilization of artificial intelligence in education has the potential to enhance the effectiveness and personalization of the learning process.

Practically, this study offers several advantages. For teachers, it provides new information and technology-based teaching strategies that can be applied to develop students' higher-order thinking skills. For schools, it can serve as a reference in developing digital-based curricula that are adaptive to changes in technology and the times. For students, it provides a more engaging and interactive learning experience while encouraging critical and creative thinking skills in solving mathematics problems.

This study is motivated by the low mathematical literacy of students at MTs Al-Ihsan in solving HOTS problems. This condition poses a serious challenge in efforts to strengthen higher-order thinking skills at the secondary education level. The issue needs to be addressed through a learning approach that can optimally facilitate students' cognitive development.

The deep learning approach was chosen for its ability to personalize learning based on an analysis of student performance data. This approach is expected to provide guidance tailored to the level of difficulty faced by students while enhancing learning effectiveness. With deep learning, students can receive adaptive instruction that meets their individual needs. The conceptual framework of this study posits that students' mathematical literacy can be improved through the integration of the deep learning approach into the learning process. Using a quantitative design, the researcher will measure the effectiveness of this approach through pre-test and post-test results on HOTS problems. The study is expected to address the main research problem and provide strategic recommendations for educators in integrating technology into mathematics instruction. Thus, the implementation of deep learning is anticipated to impact not only academic outcomes but also the comprehensive development of students' thinking skills.

Null Hypothesis (H_0): There is no significant difference in the mathematical literacy skills of students who receive learning using the deep learning model compared to students who receive conventional learning in solving HOTS problems at MTs Al-Ihsan.

Alternative Hypothesis (H_1): There is a significant difference in the mathematical literacy skills of students who receive learning using the deep learning model compared to students who receive conventional learning in solving HOTS problems at MTs Al-Ihsan.

METHOD

This study uses a quantitative approach because its main focus is to measure and analyze the effect of implementing the deep learning approach on students' literacy skills in solving HOTS problems. The quantitative approach is considered appropriate as it provides objective, systematic, and measurable results through the processing of numerical data collected from pre-test and post-test scores.

The quantitative approach emphasizes hypothesis testing, data collection using valid and reliable instruments, and statistical analysis. The primary goal of this approach is to determine the cause-and-effect relationship between the independent variable (deep learning approach) and the dependent variable (students' literacy skills).

The type of research used is a quasi-experiment with a pretest-posttest control group design. In this design, two groups are used: an experimental group and a control group. The experimental group receives

treatment through learning assisted by the deep learning approach, while the control group receives conventional learning. This design was chosen because it allows for a comparison of the treatment's effect on two relatively equivalent groups. Although full randomization is not conducted, this design still enables control over external variables that may influence the research results, thereby providing a valid representation of the treatment's effectiveness. The research subjects were two eighth-grade classes at MTs Al-Ihsan Jombang in the 2025/2026 academic year, consisting of 54 students (27 in the experimental class and 27 in the control class). The experimental class received instruction using the deep learning approach, which emphasizes mindful, meaningful, and joyful learning, while the control class was taught using conventional teaching methods.

The research instruments included a learning module, observation sheets, and pretest and posttest questions, all of which were validated by experts with a validity level of 96.5%. Data analysis was carried out using descriptive and inferential statistics through SPSS 25, including normality tests, homogeneity tests, t-tests, and N-Gain calculations.

RESULT AND DISCUSSION

Result

This study was conducted using a quasi-experimental design of the two-group pretest-posttest type. The research involved two groups of subjects: an experimental group using the deep learning approach and a control group using the conventional approach, with 27 students in class VII B and 27 students in class VII C at MTs Al-Ihsan Kalikejambon. The study was carried out from October 16–18, 2025, with the aim of examining the effect of deep learning-based instruction on students' mathematical literacy in understanding problems related to the topic of circles.

The researcher implemented the learning process according to the instructional tools designed for the deep learning approach. The study was divided into three stages: pre-implementation, implementation, and post-implementation. In the pre-implementation stage, activities included the preparation of research instruments in June 2025 and instrument validation on July 27, 2025. The implementation stage took place on October 16 and 18, 2025, at MTs Al-Ihsan Kalikejambon Tembelang Jombang, with 27 students in the control class and 27 students in the experimental class, and the material tested was circles. Finally, in the post-implementation stage, effectiveness tests were conducted, including normality tests, homogeneity tests, t-tests, and N-Gain tests. The instruments used in this study consisted of a teaching module based on the deep learning approach and assessment tests.

Discussion

The implementation of the deep learning approach was carried out in two sessions. The first session included administering the pretest and delivering the core material, while the second session involved group problem-solving activities and administering the posttest. The experimental class demonstrated active engagement and collaboration throughout the learning process, in accordance with the stages of mindful, meaningful, and joyful learning.

The researcher developed two teaching modules: the first module was used for the experimental class with the deep learning approach, and the second module was for the control class using the conventional approach.

The first module, designed by the researcher, served as a guide for conducting learning in the experimental class. This module was structured based on the deep learning approach, consisting of three main components: mindful learning, meaningful learning, and joyful learning. The module was implemented over two sessions. In the first session, students completed the pretest and received the instructional material. In the second session, students worked on problems in groups and completed the posttest.

The second module, prepared using a conventional approach, was applied in the control class. It consisted of four stages: stimulation, problem identification, discussion, and explanation. This module was also implemented in two sessions. In the first session, students completed the pretest and received the material, while in the second session, they completed the posttest. The tests were used to determine the validity of the instruments developed by the researcher for implementing the deep learning-based instruction. The research instruments consisted of pretest and posttest questions.

The research instruments in the form of pretest and posttest were validated by students, indicating that they were highly valid. The test results were obtained based on the following data:

Table 1. Pretest Validity Test

		Correlations					
		P01	P02	P03	P04	P05	TOTAL
P01	Pearson Correlation	1	.632**	.743**	.544**	.704**	.867**
	Sig. (2-tailed)		.000	.000	.003	.000	.000
	N	27	27	27	27	27	27
P02	Pearson Correlation	.632**	1	.597**	.550**	.789**	.830**
	Sig. (2-tailed)	.000		.001	.003	.000	.000
	N	27	27	27	27	27	27
P03	Pearson Correlation	.743**	.597**	1	.585**	.625**	.850**
	Sig. (2-tailed)	.000	.001		.001	.000	.000
	N	27	27	27	27	27	27
P04	Pearson Correlation	.544**	.550**	.585**	1	.667**	.786**
	Sig. (2-tailed)	.003	.003	.001		.000	.000
	N	27	27	27	27	27	27
P05	Pearson Correlation	.704**	.789**	.625**	.667**	1	.892**
	Sig. (2-tailed)	.000	.000	.000	.000		.000
	N	27	27	27	27	27	27
TOTAL	Pearson Correlation	.867**	.830**	.850**	.786**	.892**	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	
	N	27	27	27	27	27	27

** Correlation is significant at the 0.01 level (2-tailed).

The results of the pretest validity test showed that each statement item (P01 to P05) had a significant correlation with the total score, with Pearson correlation values within a strong range and a Sig. (2-tailed) value of < 0.05. This means that each item was able to measure the intended aspect and contributed to the overall construct of the questionnaire. This significant correlation indicates that the instrument used in the pretest stage was appropriate and valid for measuring the students' initial abilities or responses before the learning treatment was given.

Table 2. Posttest Validity Test

		Correlations					
		P01	P02	P03	P04	P05	TOTAL
P01	Pearson Correlation	1	.753**	.821**	.766**	.579**	.932**
	Sig. (2-tailed)		.000	.000	.000	.002	.000
	N	27	27	27	27	27	27
P02	Pearson Correlation	.753**	1	.660**	.637**	.422*	.835**
	Sig. (2-tailed)	.000		.000	.000	.028	.000
	N	27	27	27	27	27	27
P03	Pearson Correlation	.821**	.660**	1	.608**	.496**	.856**
	Sig. (2-tailed)	.000	.000		.001	.009	.000
	N	27	27	27	27	27	27
P04	Pearson Correlation	.766**	.637**	.608**	1	.597**	.852**
	Sig. (2-tailed)	.000	.000	.001		.001	.000
	N	27	27	27	27	27	27
P05	Pearson Correlation	.579**	.422*	.496**	.597**	1	.727**
	Sig. (2-tailed)	.002	.028	.009	.001		.000
	N	27	27	27	27	27	27
TOTAL	Pearson Correlation	.932**	.835**	.856**	.852**	.727**	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	
	N	27	27	27	27	27	27

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

The results of the posttest validity test also showed that all statement items had a strong correlation with the total score and a significance value of < 0.05. This indicates that after the learning treatment, the instrument continued to function well and consistently in measuring students' responses. Each item contributed stably to the overall construct, so the questionnaire instrument used in the posttest can be declared valid for measuring changes or improvements in responses after the learning process was carried out.

Table 3. Pretest Reliability Test

Reliability Statistics

Cronbach's Alpha	N of Items
.898	5

The Cronbach's Alpha value in the pretest reliability test was 0.898. This value is above the minimum reliability threshold of 0.70, indicating that the instrument can be categorized as highly reliable. This means that the five statement items used in the pretest questionnaire demonstrated good internal consistency, so the students' responses can be considered stable and not random.

Table 4. Posttest Reliability Test

Reliability Statistics

Cronbach's Alpha	N of Items
.895	5

For the posttest, the Cronbach's Alpha value was 0.895, which falls into the category of very high

reliability. Thus, the questionnaire instrument in the posttest showed good consistency in measuring students' responses after the learning process. This confirms that the instrument was not only valid but also reliable in both measurement conditions — before and after the learning treatment.

Implementation

The Deep Learning approach in the context of education does not refer to the use of artificial intelligence technology, but rather a strategy that emphasizes deep conceptual understanding rather than mere memorization of facts. In this approach, students are encouraged to learn meaningfully by connecting the material to real-life experiences and contexts, to learn joyfully through an interactive and enjoyable classroom environment, and to learn mindfully, enabling them to focus, reflect, and understand their own learning processes. By integrating these three aspects, deep learning promotes critical thinking, the development of higher-order thinking skills (HOTS), and deeper, sustainable learning motivation.

The first step taken by the researcher before starting the study was conducting observations and coordinating with the school. The researcher met with the mathematics teacher, Mrs. Daimatul Munawaroh, S.Pd., at MTs Al-Ihsan Kalikejambon to obtain permission to conduct the research. Based on this coordination, the researcher was assigned two classes as the research sample: class VII B with 27 students and class VII C with 27 students. The study was conducted from October 16 to 18, 2025.

Learning in each class was conducted over 2×2 lesson hours. In the control class, the first session took place on October 16 during the second and third periods, when the students were in a good condition to receive instruction. Activities in the first session included the students completing the pretest. The second session was held on October 17 during the fifth and sixth periods, consisting of 40 minutes of material delivery followed by discussion activities and the posttest.

In the experimental class, the first session was held on October 17 during the third and fourth periods, where students completed the pretest. The second session took place on October 18 during the first and second periods, including material delivery, group formation, quiz games, and awarding rewards to the winning groups, followed by posttest completion and a student response questionnaire.

Descriptive analysis

Descriptive analysis was conducted to provide an initial overview of students' learning outcomes before and after the treatment was applied. The data analyzed included pretest and posttest scores from both the experimental and control classes. The pretest was administered to determine students' initial ability in solving HOTS-based literacy problems before the implementation of a specific learning model. The posttest was conducted after the learning process to observe changes in students' abilities. Additionally, this analysis included calculations of the mean, minimum and maximum values, as well as data dispersion, serving as a basis for further analysis.

In the experimental class, the pretest average score indicated that most students were still in the low to moderate ability category. This suggests that students required a learning strategy capable of enhancing higher-order thinking skills and deep conceptual understanding. After the implementation of the deep learning-based instructional model, the posttest average score showed a significant improvement. This increase was not only reflected in the average scores but also in the greater variety of answers and the students' ability to explain their reasoning and problem-solving steps.

Meanwhile, in the control class using the conventional learning method, an increase in posttest scores was also observed, but it was not as pronounced as in the experimental class. The average posttest scores in the control class still fell within the low to moderate category. This indicates that, although learning had some impact on students' abilities, it was less effective in enhancing higher-order thinking skills and deep conceptual understanding compared to the deep learning-based approach.

The comparison of scores between the two classes shows a clear difference in learning effectiveness. The average N-Gain score in the experimental class was in a higher category than that of the control class, indicating that the improvement in students' abilities in the experimental class was greater. This demonstrates that the implementation of the deep learning method provides a more meaningful learning experience, where students not only receive information but also process, interpret, and apply it in problem-solving contexts.

Overall, the results of the descriptive analysis indicate that the application of the deep learning method has a positive impact on improving students' literacy skills in solving HOTS problems. Student-centered, collaborative learning processes that emphasize the development of critical, analytical, and reflective thinking skills were key factors in this improvement. Therefore, the use of the deep learning method is recommended as an alternative instructional strategy to enhance the quality of learning and student competencies at the madrasah education level.

In research, normality testing is required to ensure that the data are normally distributed. A distribution is considered normal if the significance level is greater than 0.05; conversely, if the significance level is less than 0.05, the distribution is considered non-normal. In this study, the normality of the data was tested using the Shapiro-Wilk test with the assistance of IBM SPSS Statistics 25. The data collected in this study consisted of pretest and posttest scores. The results of the normality test calculations are as follows:

Table 5. Normality Test

kelas		Tests of Normality					
		Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
hasil	pretest kontrol	.076	27	.200*	.971	27	.630
	posttest kontrol	.129	27	.200*	.964	27	.459
	pretest eksperimen	.081	27	.200*	.956	27	.294
	posttest eksperimen	.077	27	.200*	.955	27	.285

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

The table above shows the results of the normality test using the Kolmogorov–Smirnov and Shapiro–Wilk methods for both the control and experimental groups. All significance (Sig.) values for the Shapiro–Wilk test are greater than 0.05—namely 0.630, 0.459, 0.294, and 0.285—indicating that the data in all groups are normally distributed. This means that the assumption of normality required for parametric statistical tests is satisfied, and further analyses such as t-tests can be appropriately conducted using these data.

From the table, the Shapiro-Wilk significance (Sig.) values for all datasets are greater than 0.05, as follows:

Pretest (control) = 0.630

Posttest (control) = 0.459

Pretest (experimental) = 0.294

Posttest (experimental) = 0.285

Since all Sig. values > 0.05, the data are normally distributed. This means that both the control and experimental groups have pretest and posttest data that meet the normality assumption, allowing the use of parametric tests

The homogeneity test aims to determine whether the variances between the experimental and control classes are equal (homogeneous) or different. This test was conducted using Levene’s Test for Equality of Variances in SPSS. The decision criteria are as follows: if the significance value (Sig.) > 0.05, the data variances are considered homogeneous; if Sig. < 0.05, the data variances are considered not homogeneous.

Table 6. Homogeneity Test

Test of Homogeneity of Variances

		Levene Statistic	df1	df2	Sig.
Hasil	Based on Mean	3.495	1	52	.067
	Based on Median	2.548	1	52	.117
	Based on Median and with adjusted df	2.548	1	38.070	.119
	Based on trimmed mean	3.207	1	52	.079

The results of the homogeneity of variance test using Levene’s test showed a significance value (Sig.) of 0.067 based on the mean calculation, which is greater than 0.05. Therefore, it can be concluded that the variances of the two data groups are homogeneous or equal. This means there is no significant difference in variability between the control class and the experimental class. This condition meets one of the prerequisites for parametric statistical analysis, particularly for the t-test, allowing subsequent tests to be conducted without the need for corrections or alternative methods.

The results of the homogeneity test showed that the significance values for both groups were above 0.05, indicating that the variances of the two groups are homogeneous. Therefore, hypothesis testing can be conducted using the Independent Samples t-test.

The t-test is used to determine whether there is a significant difference in learning outcomes between students in the experimental class, which used the deep learning-based instructional approach for

HOTS literacy, and the control class, which used conventional learning methods. The decision criteria are as follows: if Sig. (2-tailed) < 0.05, there is a significant difference between the two groups; if Sig. (2-tailed) > 0.05, there is no significant difference.

Table 7. Uji-t (*Independent sample t-test*)

		Levene's Test for Equality of Variances				t-test for Equality of Means		95% Confidence Interval of the Difference		
Hasil		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Equal variances assumed		3.495	.067	-6.020	52	.000	-16.667	2.769	-22.222	-11.111
	Equal variances not assumed			-6.020	42.477	.000	-16.667	2.769	-22.252	-11.081

This table presents the results of the Independent Samples t-test, which was used to compare the means of two different groups. Based on Levene's test results, the significance value (Sig.) of 0.067 (> 0.05) indicates that the variances of the two groups can be considered equal; therefore, the interpretation uses the row labeled "Equal variances assumed." The t-value of -6.020 with a Sig. (2-tailed) of 0.000 (< 0.05) indicates that there is a statistically significant difference between the two groups tested. The mean difference of -16.667 shows that the first group had a lower average score than the second group, with a 95% confidence interval ranging from -22.252 to -11.081.

Table 8. Eksperimen class N-Gain Test

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
ngainscore	27	.50	1.00	.7608	.16239
ngainpersen	27	50.00	100.00	76.0812	16.23875
Valid N (listwise)	27				

This table presents descriptive statistics for two variables, namely ngainscore and ngainpersen, each with a sample size (N) of 27. The ngainscore variable has a minimum value of 0.50, a maximum value of 1.00, a mean of 0.7608, and a standard deviation of 0.16239. Meanwhile, the ngainpersen variable has a minimum value of 50.00, a maximum value of 100.00, a mean of 76.0812, and a standard deviation of 16.23875. These results indicate that, in general, the participants' gain scores were relatively high, with a moderate level of variation among individuals.

Based on the t-test results, the Sig. (2-tailed) value was found to be less than 0.05. Therefore, it can be concluded that there is a significant difference in learning outcomes between the experimental and control classes. This indicates that the learning implemented in the experimental class was more effective than that in the control class.

The calculations show that the experimental class had an average N-Gain score in the medium to high category, while the control class was in the low to medium category. This indicates that the deep learning-based instructional approach used in the experimental class was more effective in enhancing students' HOTS literacy skills compared to the conventional learning method applied in the control class.

Table 9. Control class N-Gain Test

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
ngainscore	27	-.06	.86	.4323	.24287
ngainpersen	27	-6.45	85.71	43.2342	24.28677
Valid N (listwise)	27				

This table presents the descriptive statistical results for two variables, namely ngainscore and ngainpersen, each with a sample size (N) of 27. The ngainscore variable has a minimum value of -0.06, a maximum value of 0.86, a mean of 0.4323, and a standard deviation of 0.24287. Meanwhile, the ngainpersen variable has a minimum value of -6.45, a maximum value of 85.71, a mean of 43.2342, and a standard deviation of 24.28677. The mean values, which fall within the medium range, indicate that the participants' score improvements were at a moderate level.

These findings demonstrate that the deep learning approach is effective in improving students' higher-order thinking skills and mathematical literacy. The results of this study are consistent with Pratama & Lestari (2022) and Fitriyani & Suryana (2023), who stated that adaptive learning strategies can enhance students' cognitive engagement and analytical abilities.

CONCLUSION

Based on the results and discussion of the study on “The Implementation of the Deep Learning Approach to Improve Mathematical Literacy in HOTS Problems at MTs Al-Ihsan,” several conclusions can be drawn. First, the application of the deep learning approach in mathematics instruction, which incorporates three main components—meaningful learning, joyful learning, and mindful learning—creates a more active, meaningful, and enjoyable learning environment. This learning process encourages students to think critically, reflectively, and develop a deep understanding of concepts.

Second, statistical tests show a significant difference in mathematical literacy skills between students in the experimental and control classes. Based on the t-test results, the significance value (Sig. 2-tailed) was less than 0.05, indicating that the deep learning approach has a significant effect on improving students’ mathematical literacy. Third, the N-Gain test results revealed that the experimental class showed an improvement in the medium to high category, while the control class only reached the low to medium category. This demonstrates that the deep learning approach is more effective in enhancing students’ ability to solve HOTS problems compared to conventional methods. Therefore, the deep learning approach is proven effective in improving students’ mathematical literacy, particularly in tackling Higher Order Thinking Skills (HOTS) problems that require analysis, evaluation, and creation.

The deep learning approach has been proven effective in improving students’ mathematical literacy and their ability to solve HOTS-based problems. Students who were taught through mindful, meaningful, and joyful learning demonstrated higher engagement, deeper understanding, and better reasoning skills compared to those taught through conventional methods. Statistical results (Sig. < 0.05 and N-Gain > 0.30) further support the effectiveness of this approach. Future research is recommended to apply this method to other subjects or educational levels and to integrate technology-based learning systems to support more adaptive and data-driven learning.

REFERENCES

- Hidayati, N., & Sari, R. (2023). Adaptive learning systems for mathematical improvement. *International Journal of Education and Technology*, 8(1), 45–58.
- Pratama, A., & Lestari, D. (2022). AI-based learning design for HOTS development in mathematics. *Education and Learning Research Journal*, 12(3), 134–141.
- Wahyuni, S., & Prasetyo, A. (2021). Assessing higher order thinking skills through mathematics learning. *Jurnal Pendidikan Matematika*, 15(1), 55–64.
- Zohar, A., & Dori, Y. J. (2020). Promoting higher order thinking in science and mathematics. *Educational Leadership*, 77(3), 29–35.
- Khotimah, K., & Nasrulloh, M. F. (2018). Kemampuan literasi matematika mahasiswa dengan gaya belajar sekuensial dalam menyelesaikan masalah statistika: Indonesia. *Jurnal Manajemen Pendidikan Islam Al-Idarah*, 3(2), 79–84.
- Satiti, W. S., & Khotimah, K. (2021, December 14). Soal matematika model PISA konten change and relationships untuk menunjang literasi matematis peserta didik. *Exact Papers in Compilation (EPiC)*, 3(4), 425–432.
- Khotimah, K., Nurjannah, U., & Satiti, W. S. (2024, May 5). Penerapan model eliciting activities (MEAs) dalam pembelajaran statistika untuk meningkatkan literasi matematika peserta didik kelas VIII MTsN 4 Jombang. *Al-Furqan: Jurnal Agama, Sosial, dan Budaya*, 3(3), 764–777.
- Anderson, L. W., & Krathwohl, D. R. (2010). *Kerangka landasan untuk pembelajaran, pengajaran, dan asesmen* (Terj.). Pustaka Pelajar.
- Arikunto, S. (2013). *Prosedur penelitian: Suatu pendekatan praktik*. Rineka Cipta.
- Cholik, M. (2021). Penerapan model pembelajaran untuk meningkatkan kemampuan literasi siswa. *Jurnal Pendidikan dan Pembelajaran*, 8(2), 45–54.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
- Hidayat, R., & Winata, I. (2020). Analisis kemampuan literasi sains siswa melalui soal HOTS. *Jurnal Inovasi Pendidikan Sains*, 11(1), 23–31.

- Marlina, D. (2021). Pengembangan soal Higher Order Thinking Skills dalam pembelajaran IPA di sekolah menengah. *Jurnal Evaluasi Pendidikan*, 9(3), 120–134.
- OECD. (2018). *PISA 2018 results: What students know and can do*. OECD Publishing.
- Setiawan, A., & Kurniawati, D. (2022). Penerapan deep learning dalam bidang pendidikan: Kajian dan analisis. *Jurnal Teknologi Informasi dan Pendidikan*, 15(1), 14–27.
- Sugiyono. (2019). *Metode penelitian kuantitatif, kualitatif, dan R&D*. Alfabeta.
- Zainuddin, M., & Rahmawati, S. (2021). Literasi membaca dalam pembelajaran berbasis digital. *Jurnal Pendidikan Bahasa*, 13(2), 87–98.