Hipotenusa Journal of Research Mathematics Education (2026) 17-31

Contents lists available at Universitas Muhammadiyah Lampung

Hipotenusa Journal of Research Mathematics Education

Journal homepage: https://journal.uml.ac.id/index.php/HT/index

Implementation of AI-based CTL learning model on students' mathematics learning outcomes

Yopi Ayomi Afandi[a,*,1], Noerhasmalina Alamta[a, 1], Ana Istiana[a,1], Lisdwiana Kurniati[a], Korhan Cengiz[b]

[a] *Department of Mathematics Education, Faculty of Teacher Training and Education, Universitas Muhammadiyah Pringsewu, Lampung 35373, Indonesia*

[b] *College of Information Technology, University of Fujairah, Fujairah, United Arab Emirates*

A R T I C L E I N F O

Keywords:
Artificial Intelligence;
CTL learning model;
Mathematics learning outcomes;
PjBL learning model;
Learning objective achievement criteria

A B S T R A C T

This study was based on the low learning outcomes of students at SMP Negeri 4 Gadingrejo, where only 32 out of 126 students achieved the Learning Objective Achievement Criteria (KKTP). The research method used was quantitative with an experimental design, involving two groups: the experimental class (AI-based CTL) and the control class (Project Based Learning). The instrument used was a learning outcome test consisting of 5 essay questions that had been tested for validity and reliability. Data were analyzed to measure the differences in learning outcomes between the two groups. The results showed a significant difference between the average learning outcomes of students in the experimental class and those in the control class. The average learning outcomes of students who used the AI-based CTL model were higher than those who used the PjBL model. Based on the t-test, the calculated t value was 5.13, which was greater than the t-table value of 2.00, thus the research hypothesis was accepted. The conclusion of this study is that the application of the AI-based CTL learning model is effective in improving students' learning outcomes on SPLDV material, and it is therefore recommended to consider using this model in mathematics learning.

1. Introduction

Mathematics is a universal science and plays an important role in various fields and helps develop students' thinking skills. Furthermore, mathematics also serves as a means to train students to think critically, systematically, logically, and creatively [1, 2]. According to [3], mathematics is a discipline that studies the relationships between numbers and the methods used to solve various problems involving numbers. Based on data from PISA 2022, Indonesia is ranked 67th out of 78 countries in Mathematics literacy, indicating the need for improvement in Mathematics learning [4]. This is clearly seen in SMP Negeri 4 Gadingrejo, where only 32 of 126 students in grade VIII achieved the Learning Objective Achievement Criteria (KKTP) on the Two-Variable Linear Equation System (SPLDV) material. SPLDV material has an important role in classroom learning because it is closely related to algebraic concepts and problem-solving skills [5].

* Corresponding author.
E-mail address: yopiayomi@gmail.com (Y. A. Afandi).
[1] These authors contributed equally to this work.

https://doi.org/10.36269/hjrme.v9i1.3778
Received 12 November 2025; Received in revised form 09 December 2025; Accepted 20 December 2025
Available online 6 February 2025
2723-486X/Published by Universitas Muhammadiyah Lampung.

These low learning outcomes are often caused by students' lack of understanding of the mathematical concepts taught by teachers. Previous research shows that students tend to find mathematics difficult and irrelevant to everyday life [6]. Learning models play a role in facilitating teachers in organizing classroom learning processes, from preparing teaching materials, media, and aids, to utilizing evaluation instruments to ensure learning objectives are achieved [7]. Contextual Teaching and Learning (CTL) is one approach that can increase student motivation and understanding by connecting material with real-life situations [8].

Research [9] also shows that the implementation of the CTL model is effective in improving student learning outcomes in the New Normal era. However, the integration of CTL with Artificial Intelligence (AI) technologies, such as GeoGebra, in the context of mathematics learning is still rarely studied [10]. GeoGebra is dynamic geometry software suitable for visualizing various mathematical objects because it facilitates the formation of points, lines, and various types of curves [11].

The Two-Variable Linear Equation System, commonly referred to as SPLDV in the Indonesian mathematics curriculum, is an important algebraic topic at the junior high school level [12]. This topic introduces students to mathematical situations involving two unknown quantities and two related equations. Students are expected not only to perform elimination and substitution procedures but also to identify variables, formulate mathematical models, select an appropriate solution strategy, and interpret the obtained solution. Therefore, successful learning of SPLDV requires the integration of conceptual understanding, procedural fluency, mathematical representation, and problem-solving skills [5,13,14].

In practice, many students experience difficulty when translating contextual problems into mathematical equations. They may understand the general story presented in a problem but remain unable to determine the unknown quantities, define appropriate variables, or express the relationships among those quantities symbolically. Students may also memorize elimination and substitution procedures without understanding why these procedures produce the solution of the system. This procedural orientation makes students vulnerable to errors when the form of the problem changes or when they are required to explain the meaning of the result.

SPLDV is highly relevant to contextual learning because it can represent various situations encountered in everyday life, such as determining the prices of two products, comparing the number of different objects, calculating transportation costs, or analyzing combinations of quantities. When these situations are presented meaningfully, students can recognize that mathematical equations are not merely abstract symbols but representations of relationships found in real conditions. Contextual problems can therefore serve as a bridge between students' informal understanding and formal algebraic reasoning [6,8].

The Contextual Teaching and Learning model emphasizes the relationship between academic content and students' real-life experiences. Through contextual learning, students are encouraged to construct knowledge, investigate problems, ask questions, collaborate, observe mathematical models, reflect on their learning, and complete meaningful assessments. In mathematics instruction, these activities can help students understand why a mathematical procedure is needed and how the resulting answer relates to the original problem. Previous studies have shown that CTL can support mathematical problem solving and improve students' learning outcomes [6,9].

However, the effectiveness of CTL depends on how contextual situations are connected to formal mathematical concepts. Merely including a story problem does not automatically make learning contextual. Students need guidance to identify important information, define variables, formulate equations, solve the system, and return the solution to its original context. Without this connection, students may understand the situation but fail to develop the intended algebraic knowledge. The teacher therefore plays an important role in providing scaffolding and directing students from concrete situations toward symbolic and graphical representations.

The use of multiple mathematical representations is especially important in learning SPLDV. A system of equations may be expressed through verbal descriptions, algebraic symbols, tables, and graphs. Each representation provides different information about the same mathematical relationship. Symbolic forms allow students to perform algebraic operations, whereas graphical forms show the intersection of two linear equations visually. Students who can connect these representations are more likely to understand that the solution of an SPLDV is not only an ordered pair obtained through calculation but also the point satisfying both equations [13,14].

Nevertheless, drawing graphs manually can create additional difficulties. Students must determine coordinate points, select an appropriate scale, plot the points accurately, and construct straight lines. Errors in one of these stages may lead to incorrect conclusions about the solution. The manual process may also require considerable instructional time, limiting opportunities

for students to compare equations and explore changes in their graphical forms. Digital mathematical tools can reduce these technical barriers and allow students to focus more on interpretation and conceptual relationships.

GeoGebra provides dynamic connections among algebraic, numerical, and graphical representations. In SPLDV learning, students can enter two equations and immediately observe the corresponding straight lines and their intersection point. They can also modify coefficients or constants and examine how these changes affect the position and intersection of the lines. These functions make it possible for students to explore systems with one solution, no solution, or infinitely many solutions in a more visual and interactive manner [11,15].

GeoGebra can also function as a verification tool. After students solve an SPLDV problem using elimination or substitution, they can compare the calculated ordered pair with the intersection displayed by the application. When the two results are inconsistent, students can review their calculations, reconsider the equations they formulated, or identify errors in entering the expressions. This immediate visual response can support formative learning because students receive information that helps them evaluate and revise their reasoning.

The integration of GeoGebra into CTL creates complementary instructional functions. CTL provides meaningful problems and a learning sequence that connects mathematics with everyday situations, while GeoGebra supports visualization, exploration, and verification. Students may begin by analyzing a contextual problem, defining variables, and constructing two equations. They can then use GeoGebra to examine the graphical relationship between the equations before interpreting the intersection point according to the original problem. Through this sequence, technology supports mathematical reasoning rather than replacing it.

Technology integration should nevertheless be accompanied by appropriate teacher guidance. Students' familiarity with smartphones or computers does not necessarily mean that they can use mathematical software effectively. They need clear instructions for entering equations, reading coordinate values, interpreting graphs, and relating digital outputs to algebraic procedures. Teachers must also ask questions that direct students toward mathematical meaning, such as why two lines intersect, what the intersection represents, and how graphical and algebraic solutions are related.

In this study, the expression "AI-based CTL" is operationally used to describe a technology-assisted contextual learning process supported by dynamic digital computation and visualization through GeoGebra. GeoGebra itself is primarily dynamic mathematics software rather than a conventional generative artificial intelligence system. Therefore, the technological contribution examined in this research is specifically associated with its ability to generate representations, provide immediate visual responses, and support students' exploration of mathematical relationships. This clarification is necessary to avoid interpreting the role of AI beyond the actual technological features used during the intervention.

Project Based Learning, which serves as the comparison model in this study, is also a student-centered approach that can promote collaboration, creativity, responsibility, and active participation. However, PjBL generally organizes learning around the completion of a project or product, whereas CTL emphasizes continuous connections between academic content and meaningful situations. For a focused algebraic topic such as SPLDV, CTL supported by dynamic visualization may provide more direct opportunities for students to analyze several contextual problems, formulate equations, compare solution strategies, and verify their answers within the available instructional time.

This comparison does not imply that PjBL is generally less effective than CTL. The effectiveness of a learning model depends on the objectives, characteristics of the material, duration of instruction, teacher competence, student readiness, and form of assessment. PjBL may be highly appropriate for learning objectives that emphasize extended investigation and product development. By contrast, technology-assisted CTL may be particularly relevant when the objective is to strengthen students' understanding of relationships between contextual situations, algebraic equations, and graphical representations.

Previous studies have separately reported the benefits of CTL, GeoGebra, digital learning media, and innovative mathematics learning models [8–11]. However, research directly comparing a CTL model supported by GeoGebra with PjBL in junior high school SPLDV learning remains limited. In addition, several studies discuss educational technology without clearly explaining how the tool is integrated into the pedagogical stages. In the present study, GeoGebra is positioned not merely as a presentation medium but as a tool for exploring, visualizing, and verifying the mathematical models constructed from contextual problems.

The novelty of this study lies in the integration of contextual SPLDV problems, CTL learning stages, and dynamic visualization through GeoGebra, followed by a direct comparison with the PjBL model. The study focuses on whether this integrated learning process produces different mathematics learning outcomes among Grade VIII students. It also provides classroom-based evidence from SMP Negeri 4 Gadingrejo, where the low proportion

of students meeting the established learning criteria indicates a practical need for instructional improvement.

Based on these considerations, an investigation is required to determine whether students taught using AI-based CTL supported by GeoGebra obtain significantly different learning outcomes from students taught using PjBL. The comparison is expected to provide empirical information regarding the suitability of technology-supported contextual learning for SPLDV instruction. It may also serve as a reference for mathematics teachers seeking to combine contextual problems, multiple representations, and digital visualization in order to improve students' engagement and mathematical understanding.

This study aims to provide an alternative to improve junior high school students' Mathematics learning outcomes, especially in SPLDV material. The proposed hypothesis is: "There is a significant difference between the average learning outcomes of students who use the AI-based CTL Model and students who use the Project Based Learning (PjBL) Model."

The purpose of this study was to analyze the effect of implementing the AI-based CTL learning model on junior high school students' learning outcomes in SPLDV material. Thus, it is hoped that this research can contribute to the development of more effective and relevant learning methods.

2. Method

2.1. Research design

This study employed a quantitative method with a comparative experimental design. The objective was to analyze the differences in the influence of the AI-based CTL Learning Model on student learning outcomes in the SPLDV material compared to the Project Based Learning (PjBL) model. The design involved two groups: an experimental class using the AI-based CTL model (GeoGebra) and a control class using the PjBL model.

2.2.Research Setting and Participants

The research was conducted at SMP Negeri 4 Gadingrejo, Lampung, during the 2024/2025 academic year. The population consisted of 126 students divided into four classes (VIII A, B, C, and D).

The participants were selected using cluster random sampling. According to [12], this technique is used when the population consists of groups of individuals or clusters rather than single individuals. Based on this, Class VIII C was designated as the experimental class applying the AI-based CTL model, and Class VIII D as the control class using the PjBL Model [10, 13, 14].

2.3.Development Procedure

The procedure for executing this research followed systematic stages:

1. Preparation Stage
   The researcher prepared the learning outcome test instrument in the form of essay questions designed to measure in-depth understanding of SPLDV concepts. The instrument was then tested for validity and reliability before use.
2. Implementation Stage
   The experimental class underwent learning using the AI-based CTL model with the GeoGebra application, while the control class used the PjBL model.
3. Evaluation Stage
   After the learning process, both classes were given a post-test to measure student learning outcomes and analyze the differences.

2.4.Instruments and Data Collection

The instrument used was a learning outcome test consisting of 5 essay questions. Before the research, the instrument underwent a trial (*try out*) to determine:

1. Validity
   The content validity test is conducted using a validation sheet and test equipment validated by mathematics teachers and lecturers [15]. Item validity was calculated using the Product Moment correlation formula.
2. Difficulty Level
   Questions that are too easy or too difficult have poor discriminatory power because they fail to differentiate student abilities [16–18]. Therefore, a question is considered good if it has a moderate difficulty index that meets the criteria [19–21].
3. Reliability
   Tested using Cronbach's Alpha formula to measure the consistency of the test items.

2.5.Data Analysis

The data analysis techniques included:

1. Prerequisite Tests: Normality tests (using Chi-Square) to check data distribution and homogeneity tests (using F-test). If $F_{count} < F_{table}$, it means the data is homogeneous [22].
2. Hypothesis Testing: Using the *t-test* (both two-tailed and one-tailed) with a significance level of 5% to determine the significant difference in average learning outcomes between the experimental and control classes.

3. Results and discussion

3.1. Analysis Stage

The analysis stage involved evaluating the quality of the research instrument through validity and reliability testing on 126 students.

a) Validity Test
The content validity was first assessed by experts. Then, item validity was calculated. The results showed that all 6 question items were valid because the calculated t value was greater than the $t_{table}(0.374)$.

Table 1
Results of the Validity Test of the Test Instrument

No. Item Question	Price T count	Price T table	Decision
1	4,915	0,374	Valid
2	4,622	0,374	Valid
3	7,246	0,374	Valid
4	4,245	0,374	Valid
5	10,989	0,374	Valid
6	7,907	0,374	Valid

(Source: Excel 2007 Calculation Results).

b) Reliability Test
The reliability of the instrument was tested using Cronbach's Alpha. The result showed a value of 0.83, which is greater than 0.60, indicating that the instrument is reliable.

Table 2
Results of the Reliability Test of the Test Instrument

Reference Value	Cronbach Alpha	Conclusion
0,60	0,83	RELIABLE

Since α > 0.60, the instrument was declared reliable.

The validity and reliability results indicate that the instrument was sufficiently appropriate for measuring students' learning outcomes in SPLDV. The validity test was intended to determine whether each item measured the mathematical competencies targeted in the study. Because the reported test statistics for all items exceeded the criterion value, the items were retained for use in the experimental stage. This finding suggests that the questions were capable of representing the intended content and distinguishing students according to their performance on the measured competencies.

Instrument validity is particularly important in SPLDV assessment because the topic involves several interconnected abilities. Students must understand the problem context, define variables, formulate equations, apply an appropriate solution procedure, and interpret the solution. Therefore, the assessment should not measure numerical calculation alone. The use of essay questions allowed students to demonstrate their reasoning process and provided information about how they constructed and solved the system of equations.

The use of expert judgment before empirical testing also strengthened the quality of the instrument. Mathematics teachers and lecturers could evaluate the correspondence between the questions, learning indicators, curriculum content, and cognitive level expected from Grade VIII students. This process reduced the possibility that an item would be statistically acceptable but conceptually inappropriate. Content validation and empirical validation therefore served complementary purposes in establishing the quality of the assessment.

The reliability coefficient of 0.83 indicates a high level of internal consistency among the items. This result means that the items generally measured related aspects of students' SPLDV competence. A reliable instrument is essential in comparative research because inconsistent measurement can obscure the actual differences between instructional groups. The obtained coefficient supports the use of the test for comparing the learning outcomes of students taught using AI-based CTL and those taught using PjBL.

Nevertheless, a high reliability coefficient does not mean that all items measured exactly the same skill. SPLDV learning outcomes may include conceptual understanding, procedural fluency, mathematical modeling, graphical interpretation, and problem solving. Variation among the items is therefore still necessary to represent these different dimensions. An instrument consisting of identical or overly similar questions may produce high consistency but provide a narrow description of students' competence.

Because the instrument used essay questions, scoring consistency should also be considered. Essay responses may contain different solution strategies even when students reach the same answer. A clear scoring rubric is therefore required to determine how points are awarded for defining variables, constructing equations, selecting a method, performing calculations, and interpreting the result. The use of a detailed rubric can reduce subjectivity and improve the consistency of scoring across students.

Future studies may strengthen the assessment procedure by involving more than one scorer and examining inter-rater agreement. Independent scoring by two assessors would provide evidence that the students' scores were not strongly influenced by individual judgment. Disagreements between scorers could be discussed using predetermined criteria before the final scores are established.

It is also important to ensure consistency in reporting the number of test items. The Method

section states that the learning-outcome test consisted of five essay questions, whereas the validity and difficulty analyses present six items. This discrepancy should be corrected because it affects the transparency and reproducibility of the research. The manuscript should clarify whether six items were initially tested and five were selected for the post-test, or whether all six items were ultimately administered.

Overall, the instrument analysis provided an adequate basis for proceeding to the comparative test. The combined evidence from expert validation, empirical validity, and reliability testing suggests that the learning-outcome scores were obtained using an instrument with acceptable measurement quality. However, clearer reporting of the item-selection process and scoring procedure would further strengthen the credibility of the findings.

3.2.Design Stage

In this stage, the design of the test items was analyzed for their difficulty level. The analysis showed that question number 1 was categorized as "Easy," while questions 2, 3, 4, 5, and 6 were categorized as "Medium". This confirms the instrument meets the design criteria for moderate difficulty.

Table 3
Results of Test the Level of Difficulty of Question

No. Question Item	P_{count} (Difficulty Level)	Interpretation of Difficulty Index
1	0,76	Easy
2	0,64	Currently
3	0,58	Currently
4	0,52	Currently
5	0,59	Currently
6	0,61	Currently

The difficulty-index results show that the test was dominated by items in the medium category. Question 1 had a difficulty index of 0.76 and was categorized as easy, whereas the remaining items had indices ranging from 0.52 to 0.64 and were categorized as medium. This distribution indicates that most questions were neither too easy nor too difficult for the participating students.

Items with moderate difficulty are generally useful in comparative studies because they provide sufficient opportunities for students with different levels of ability to demonstrate their competence. Extremely easy items may be answered correctly by almost all students and therefore provide limited information for distinguishing between groups. Conversely, extremely difficult items may produce uniformly low scores and fail to reveal meaningful differences in students' understanding.

The inclusion of one easy item may also have a useful instructional and measurement function. An easier introductory question can help students become familiar with the format of the assessment and reduce excessive anxiety at the beginning of the test. It may also measure whether students possess the basic prerequisite knowledge needed before completing more demanding SPLDV problems.

However, the absence of difficult items should be considered when interpreting the test results. A test containing only easy and medium questions may adequately measure basic and intermediate achievement but may be less effective in identifying students with advanced reasoning ability. Difficult questions can provide information about whether students are able to solve unfamiliar problems, integrate several representations, or justify complex mathematical decisions.

Future instruments may therefore include a more balanced combination of easy, medium, and difficult items. The difficult items should not be difficult because of complicated language or excessive calculations. Instead, their difficulty should arise from higher cognitive demands, such as constructing equations from an unfamiliar context, comparing alternative solution strategies, interpreting graphical conditions, or explaining why a system has no solution or infinitely many solutions.

The difficulty index should also be interpreted together with item validity and discriminatory power. An item may have moderate difficulty but still be ineffective when it cannot distinguish between students with high and low competence. Conversely, an item may appear difficult because its instructions are unclear rather than because it measures advanced mathematical reasoning. Examining these indicators together provides a more comprehensive evaluation of item quality.

For essay questions, the difficulty index may also be influenced by partial-credit scoring. Students who construct the equations correctly but make an arithmetic error may still receive part of the maximum score. This characteristic is beneficial because it distinguishes conceptual errors from procedural errors. The manuscript should therefore explain whether the difficulty indices were calculated using total item scores or dichotomous correct–incorrect classifications.

The predominance of medium-difficulty items suggests that the instrument was generally appropriate for the Grade VIII participants. It reduced the likelihood of severe floor or ceiling effects and provided a reasonable basis for comparing the two instructional groups. Nevertheless, adding one or two higher-order items in future research would enable a more detailed evaluation of students' advanced SPLDV reasoning.

3.3.Development Stage

The development stage in this context refers to establishing the statistical model's validity through prerequisite testing.

a) Normality Test
Experimental Class: $X^2_{count}(6.93) < X^2_{table}(7.815)$, meaning the data is normally distributed.

Table 4
Experimental Class Normality Test

X² count	X²table
6,93	7,815

Control Class: $X^2_{count}(6.80) < X^2_{table}(7.815)$, meaning the data is normally distributed.

Table 5
Control Class Normality Test

X²count	X²table
6,80	7,815

c) Homogeneity Test
The homogeneity was tested by comparing variances. The result showed F_{count} $(0.567) < F_{table}(1.840)$, indicating that the data from both samples were homogeneous [22].

The normality-test results show that the calculated Chi-Square values for the experimental and control classes were below the critical value of 7.815. The experimental class obtained a value of 6.93, while the control class obtained a value of 6.80. Based on the criteria used in the study, the learning-outcome data from both groups were therefore considered normally distributed.

The fulfillment of the normality assumption supports the use of a parametric t-test. Parametric testing assumes that the distributions of the analyzed scores do not deviate substantially from normality. When this assumption is reasonably satisfied, the comparison of group means can be conducted using procedures that provide efficient statistical estimates.

The relatively similar Chi-Square values obtained by the two classes also indicate that neither group showed a markedly different distributional pattern according to the selected test. This result is relevant because severe skewness or extreme concentration of scores in one group could affect the interpretation of the difference between the means.

Nevertheless, normality should not be interpreted as evidence that the two classes had equivalent ability before the intervention. A normal distribution only describes the shape of the score distribution. It does not demonstrate that the groups had equal initial achievement, similar prior knowledge, or identical learning characteristics.

A graphical examination, such as a histogram or normal probability plot, could complement the Chi-Square test. Statistical tests of normality may be affected by sample size, whereas graphical inspection can reveal skewness, outliers, and unusual score patterns. Reporting both statistical and graphical evidence would provide a more complete description of the data distribution.

The presence of outliers should also be examined because a small number of extremely high or low scores may influence the group mean and t-test result. The manuscript does not currently report whether outliers were detected or how they were handled. Future studies should describe the criteria used to identify unusual observations and explain whether all scores were retained in the analysis.

Table 6
Homogeneity Test Results

F_{count}	F_{table}
0,567	1,840

The homogeneity test was conducted to determine whether the variability of learning-outcome scores was comparable between the experimental and control classes. The result was interpreted as indicating homogeneous variance because the calculated F value was below the critical F value. This finding fulfilled another prerequisite for the independent-samples t-test used in the study.

Homogeneity of variance means that the spread of scores in the two groups was sufficiently similar for the selected statistical comparison. This assumption is relevant because a substantial difference in variance may affect the standard error used to evaluate the difference between group means. The homogeneous result therefore supports the technical appropriateness of the hypothesis-testing procedure.

However, the exact calculated F value and critical F value must be presented in Table 6. Without these values, readers cannot independently verify the conclusion that the variances were homogeneous. The table should also report the degrees of freedom and the significance level used in determining the critical value.

Homogeneous variance should not be interpreted as proof that the groups were equivalent before instruction. Two groups can have similar score variability while having different initial means. Because the study appears to rely primarily on post-test scores, the absence of reported pretest data limits the ability to determine whether the classes had comparable SPLDV competence before the intervention.

Cluster random sampling reduces selection subjectivity by selecting intact classes rather than individual students. Nevertheless, assigning intact classes does not always guarantee complete equivalence between groups. The classes may differ in prior achievement, classroom interaction,

motivation, or familiarity with digital applications. These possible differences should be acknowledged when interpreting the post-test comparison.

Future research can address this issue by administering a pretest and analyzing either gain scores or post-test scores adjusted for initial achievement. A pretest–post-test control-group design would provide stronger evidence that differences observed after instruction were associated with the learning intervention rather than with pre-existing differences between the classes.

3.4.Implementation Stage

This stage presents the results of the actual implementation of the learning models, measured through hypothesis testing.

a) Two-tailed t-test
   The test determined if there was a difference between the two models. The result was T_{count} (5.13) $< T_{table}$(2.00), meaning H_0 was rejected. This confirms a significant difference in learning outcomes between the AI-based CTL model and the PjBL model.
b) One-tailed t-test
   This test determined which model was better. The result T_{count} (5.13) $< T_{table}$(2.00) showed that the average learning outcomes of the AI-based CTL class were significantly higher than the PjBL class.

The reported value of T_{count} (5.13), which exceeded T_{table}(2.00), indicates that the difference between the average learning outcomes of the two classes was statistically significant at the criterion used in the study. Accordingly, the null hypothesis stating that there was no difference between the two group means was rejected.

The two-tailed test establishes that the two instructional conditions produced different average outcomes, but it does not by itself identify which group obtained the higher mean. The direction of the difference must be determined by examining the descriptive statistics and the subsequent directional test. Therefore, the means, standard deviations, and number of participants in both classes should be displayed in a separate table.

Reporting descriptive statistics is important because statistical significance depends partly on sample size and score variability. Readers need to know how large the observed difference was in the original scoring units. For example, a statistically significant result may represent either a substantial improvement or a relatively small difference detected with sufficiently stable data.

The manuscript should also report the degrees of freedom and exact probability value associated with the t statistic. Reporting only the comparison between t_{count} and t_{table} provides a basic decision but does not fully meet contemporary statistical-reporting practices. An exact p value allows readers to evaluate the strength of the evidence against the null hypothesis.

A confidence interval for the difference between means would further improve the interpretation. The interval would indicate the range of plausible values for the average advantage associated with the AI-based CTL condition. A narrow confidence interval would suggest a more precise estimate, whereas a wide interval would indicate greater uncertainty.

In addition to statistical significance, an effect-size measure should be calculated. Cohen's d, Hedges' g, or another appropriate standardized measure would show the magnitude of the difference relative to score variability. Effect size is essential for determining whether the difference is educationally meaningful rather than merely statistically detectable.

These additional statistics would make the conclusion more informative. Instead of stating only that one model was more effective, the study could explain how large the difference was, how precise the estimate was, and whether the magnitude was small, moderate, or large. Such reporting would facilitate comparison with other studies examining CTL, GeoGebra, PjBL, or mathematics learning outcomes.

The one-tailed test provided evidence that the direction of the observed difference favored the AI-based CTL class. Students taught through contextual learning supported by GeoGebra obtained a higher average learning outcome than those taught through PjBL under the conditions of this study. This result supports the directional research hypothesis.

A one-tailed test is appropriate when the directional hypothesis has been established before the data are analyzed and when a result in the opposite direction would not support the theoretical expectation. The manuscript should clarify that the hypothesis predicting higher achievement in the AI-based CTL class was formulated before the statistical analysis. This clarification is needed because selecting a one-tailed test after observing the group means can increase the risk of a biased conclusion.

The higher learning outcomes in the experimental class may be associated with the alignment between CTL activities and the characteristics of SPLDV. The topic requires students to connect verbal situations with algebraic equations and interpret the resulting solution. Contextual problems provide a meaningful starting point for this process by presenting unknown quantities and mathematical relationships through situations familiar to students [6,8].

Within the CTL process, students may have been encouraged to identify relevant information, determine variables, formulate equations, and discuss possible solution strategies. These activities require active processing of mathematical information rather than passive reception of a completed procedure. Active engagement may strengthen students' understanding of how SPLDV models are constructed and why particular solution methods are used.

GeoGebra may have strengthened this process by presenting algebraic and graphical representations dynamically [11,23]. Students could observe that each

linear equation corresponds to a straight line and that the solution of the system is represented by the intersection point. This visualization provides conceptual meaning for the ordered pair obtained through elimination or substitution.

The opportunity to compare algebraic calculations with graphical results may also have functioned as formative feedback. When the calculated answer did not correspond to the displayed intersection, students could review the equations, examine their arithmetic operations, or reconsider their interpretation. This immediate response may have reduced the persistence of errors during the learning process.

However, the difference cannot be attributed to GeoGebra alone. The intervention combined the CTL model, contextual problems, teacher guidance, classroom interaction, and digital visualization. The observed learning outcomes should therefore be interpreted as the result of an integrated instructional process rather than the isolated effect of one application.

Similarly, the result should not be interpreted as evidence that PjBL is generally ineffective. PjBL is designed to support investigation, collaboration, responsibility, creativity, and the production of meaningful projects. Its effectiveness may be more apparent when learning outcomes include project quality, teamwork, communication, or long-term problem solving rather than achievement on a focused SPLDV test.

The comparative advantage of AI-based CTL in this study may reflect the close alignment between the intervention and the assessed competencies. The post-test measured students' ability to solve SPLDV essay questions. CTL supported by GeoGebra provided direct and repeated engagement with contextual SPLDV problems, equation construction, and solution verification. PjBL may have allocated more classroom time to project organization and product completion.

This interpretation emphasizes the importance of alignment among learning objectives, instructional activities, and assessment. A learning model is more likely to produce high test performance when the activities provide sustained practice in the competencies evaluated by the test. Therefore, the results should be discussed within the specific context of SPLDV learning outcomes rather than used to establish a universal ranking of the two models.

3.5.Evaluation Stage

The results showed that the AI-based CTL Learning Model is more effective than the PjBL Model. CTL encourages the brain to create patterns that embody meaning and connects academic content to the context of students' daily lives. This is supported by the use of GeoGebra, a mathematics application capable of presenting mathematical content such as geometry, algebra, statistics, and calculus applications [23].

Research conducted by [24] also shows that using the CTL model can help students improve their learning outcomes. The novelty of this research is that the AI-based CTL model can improve student learning outcomes, make students enthusiastic, actively answer questions, be responsible for completing assignments, and be able to understand SPLDV material through examples, contextual illustrations, and various learning resources [10, 18–21, 25–27].

The findings are consistent with previous studies indicating that CTL can support mathematics learning outcomes and problem-solving ability [6,9,24]. Contextual learning provides students with opportunities to connect formal mathematical content with situations that possess recognizable meaning. This connection may reduce the perception that mathematics consists only of abstract rules and encourage students to understand the purpose of the procedures they perform.

In SPLDV learning, contextualization has a particularly important function because students must formulate equations from relationships between quantities. A problem involving product prices, ticket combinations, age relationships, or the number of two types of objects can be represented as two linear equations. When students understand the situation, the equations are more likely to be perceived as meaningful representations rather than arbitrary collections of symbols.

The experimental learning process may also have supported mathematical modeling. Modeling begins when students identify the important quantities in a problem, define the unknown values, and formulate relationships in mathematical language. After solving the equations, students must interpret the ordered pair within the original situation. These stages require students to move repeatedly between contextual and mathematical representations.

The ability to move between representations is a fundamental component of mathematical understanding [13,14]. Students who rely exclusively on symbolic manipulation may obtain an answer but remain unable to explain its graphical or contextual meaning. Conversely, students who understand the situation informally may still experience difficulty expressing it through equations. The integration of CTL and GeoGebra may help connect verbal, symbolic, numerical, and graphical forms.

GeoGebra makes the relationship between the two equations visible. Each equation is represented as a line, and the common solution is displayed as the intersection of the two lines. This representation can help students understand why an ordered pair must satisfy both equations simultaneously. It also illustrates that the solution is a relationship shared by the two mathematical conditions rather than an isolated result produced by a formula.

Dynamic visualization may be especially useful when discussing different types of systems. Students can modify the coefficients and observe whether the lines intersect, remain parallel, or coincide. These visual patterns correspond to systems with one solution, no solution, or infinitely many solutions. Exploring these cases can deepen conceptual understanding beyond the

routine application of elimination and substitution.

The use of GeoGebra may also reduce the technical burden associated with manual graph construction. Drawing graphs manually requires students to calculate coordinate points, select a scale, plot the points, and draw lines accurately. Errors in these steps can obscure the mathematical relationship being studied. Digital graphing allows students to focus more directly on interpreting how changes in equations affect the resulting lines.

Nevertheless, manual and digital methods should not be treated as mutually exclusive. Manual graphing can help students understand how coordinate pairs are generated and how a linear equation is represented geometrically. GeoGebra can then extend this understanding by allowing faster comparison and exploration. A balanced learning sequence may begin with manual construction and continue with dynamic digital investigation.

The reported improvement may also be related to increased opportunities for feedback. In conventional exercises, students may complete several calculations before learning that the original equations were formulated incorrectly. GeoGebra provides an immediate graphical response to the equations entered. Unexpected graphs or intersection points can signal that a mistake has occurred and prompt students to reconsider their work.

However, immediate feedback is educationally valuable only when students are encouraged to analyze it. Students may otherwise respond by repeatedly changing values until the expected graph appears without understanding the reason. Teacher questioning remains necessary to direct students toward interpretation. Questions such as “What does this intersection represent?”, “Why are these lines parallel?”, and “Which part of the equation caused the graph to change?” can transform visual output into mathematical reasoning.

The teacher’s role is therefore central to the effectiveness of the intervention. Technology does not independently generate meaningful learning. Teachers must select appropriate contextual problems, organize the learning sequence, facilitate discussion, monitor the use of the software, and help students connect the digital output with formal mathematical concepts.

CTL also encourages students to become active participants in constructing knowledge. Instead of receiving every equation and procedure directly, students may investigate the given situation, propose mathematical relationships, compare strategies, and explain their reasoning. These activities can reveal misconceptions that might remain hidden in a lecture-centered classroom.

Questioning is an important CTL component because it stimulates explanation and provides diagnostic information. Students who are asked why they selected a particular variable or why they used elimination rather than substitution must articulate their understanding. Their responses help teachers determine whether the solution reflects genuine reasoning or memorized steps.

Learning communities may also contribute to improved outcomes. Discussion with classmates exposes students to alternative representations and solution methods. One student may solve a system using substitution, while another may use elimination or a graph. Comparing these strategies can help students recognize that different methods represent the same underlying mathematical relationship.

Collaborative discussion should nevertheless be structured carefully. In group activities, students with stronger prior knowledge may dominate the process while other members copy the solution. Teachers should assign clear responsibilities and require each student to explain part of the reasoning. Individual assessment remains important for determining whether every student has developed the intended competence.

Reflection is another mechanism that may distinguish meaningful CTL from routine contextual exercises. At the end of the activity, students can summarize the relationship among the problem context, equations, solution procedure, graph, and interpreted answer. Reflection encourages them to organize the knowledge gained during exploration and connect it with previously learned concepts.

The observed results can also be interpreted through the distinction between conceptual understanding and procedural fluency. Procedural fluency enables students to perform elimination and substitution accurately. Conceptual understanding enables them to explain why the procedures preserve the solution and how the resulting ordered pair relates to both equations. Effective SPLDV learning requires both dimensions.

GeoGebra-assisted CTL may support procedural learning by allowing students to check their calculations and conceptual learning by visualizing the graphical meaning of the solution. When these functions are integrated effectively, students are less likely to regard algebraic procedures as disconnected rules. This integration may explain why the experimental class achieved higher post-test outcomes.

The findings may also be related to student motivation. Mathematics is often perceived as difficult when students cannot identify its relevance or visualize its abstract relationships. Contextual examples can demonstrate the usefulness of SPLDV, while digital visualization may make the learning process more interactive. Together, these features may increase students’ willingness to participate and persist in completing problems.

However, motivation was not measured using a validated instrument in this study. Statements that students became more enthusiastic, active, or responsible should therefore be presented as classroom observations rather than as statistically established outcomes. Future research should use a motivation questionnaire, observation rubric, or interview protocol when drawing conclusions about affective changes.

Similarly, increased conceptual understanding should be supported by appropriate evidence. The post-test

scores demonstrate differences in mathematics learning outcomes, but they may not reveal which specific concepts improved. An item-by-item analysis could identify whether the experimental class performed particularly well in mathematical modeling, elimination, substitution, graphical interpretation, or contextual interpretation.

Analyzing students' written solutions would provide additional insight. Researchers could categorize common errors, such as defining variables incorrectly, constructing only one equation, applying an invalid operation, making arithmetic errors, or failing to interpret the final answer. A comparison of error patterns between the classes would help explain how the instructional models influenced learning.

Qualitative data could also strengthen the discussion. Interviews with selected students could explore how they experienced contextual problems, which GeoGebra features were helpful, and what difficulties remained. Classroom observations could document how students interacted with the software, discussed solutions, and responded to teacher questions. These data would explain the learning process behind the numerical difference.

The study's comparison with PjBL should also consider implementation fidelity. The effectiveness of each learning model depends on whether its defining stages were applied consistently. CTL should include meaningful contextual connections, inquiry, questioning, modeling, reflection, and authentic assessment. PjBL should involve a driving problem, planning, investigation, product development, monitoring, and evaluation.

The manuscript currently provides limited information about the duration and detailed activities of each condition. Readers need to know how many meetings were conducted, how much instructional time each group received, which teacher delivered the lessons, and how the implementation of the models was monitored. Without this information, it is difficult to determine whether the differences arose from the models themselves or from variations in teaching intensity.

Teacher effects should also be considered. When different teachers teach the two classes, differences in experience, explanation, classroom management, and familiarity with technology may influence student outcomes. When the same teacher teaches both classes, there remains a possibility of differential enthusiasm or greater preparation for the experimental condition. Future studies should describe and control these factors.

The post-test-only comparison represents another limitation. Although cluster random sampling was used, intact classes may differ before instruction. A pretest would provide empirical evidence of their initial equivalence. It would also allow the researchers to calculate learning gains and determine whether students with different initial abilities benefited similarly from the intervention.

A delayed post-test would be useful for examining retention. Digital and contextual activities may produce strong immediate performance because the learning experience is recent and engaging. However, the long-term value of the model depends on whether students retain their understanding and can apply it to new problems after several weeks.

Transfer should also be investigated. Students may learn to solve problems similar to those presented during instruction without being able to apply SPLDV to unfamiliar contexts. Future assessments should include near-transfer and far-transfer problems. Near-transfer items may change numerical values or surface features, while far-transfer items may present the same algebraic structure through substantially different situations.

The terminology "AI-based CTL" requires careful interpretation. GeoGebra is primarily dynamic mathematics software that generates algebraic and graphical representations based on user input. Unless additional adaptive, predictive, or machine-learning features were used, the intervention may be more accurately described as technology-assisted or GeoGebra-assisted CTL. The manuscript should clearly explain which specific characteristics justify the AI designation.

This clarification does not reduce the instructional value of the intervention. Dynamic computation and visualization can provide powerful support for mathematical exploration even when they do not constitute artificial intelligence in a strict technical sense. Precise terminology strengthens the scientific contribution by ensuring that readers understand exactly what technology was implemented.

The study also has practical implications for mathematics teachers. GeoGebra-assisted contextual learning can be used when introducing SPLDV, comparing solution methods, discussing graphical interpretations, or verifying students' calculations. The application may be projected by the teacher, used in a computer laboratory, or operated by students through available devices.

Implementation should consider technological access and digital competence. Students may have different levels of familiarity with GeoGebra, and not all schools have equal access to computers, smartphones, or stable internet connections. Teachers should provide an introductory tutorial and prepare alternative activities when technical problems occur.

Lesson design should prevent technology from replacing essential mathematical reasoning. Students should formulate equations before entering them into GeoGebra, predict the form of the graph before viewing it, and explain the meaning of the intersection after it is displayed. These requirements ensure that the software supports rather than bypasses reasoning.

Teachers can also integrate written work with digital exploration. Students may first solve the system manually, verify it graphically, and then write a reflection comparing the methods. This combination provides evidence of procedural fluency, conceptual understanding, and representational competence.

From a curriculum perspective, the findings support the use of contextual and technology-supported tasks in algebra instruction. Mathematical literacy requires students to formulate, employ, and interpret mathematics in different situations. SPLDV provides an appropriate context for developing these processes because students must move from real problems to equations and return from mathematical solutions to contextual conclusions.

The findings are consistent with research showing positive effects of CTL on mathematics learning [6,9,24] and with studies emphasizing the role of digital and visual representations [11,13,14,23]. The present study extends this literature by directly comparing GeoGebra-supported CTL with PjBL in SPLDV instruction at the junior high school level.

The contribution of the study lies not merely in demonstrating a statistical difference but in illustrating a possible instructional alignment. Contextual problems provide meaning, CTL organizes student-centered activities, and GeoGebra connects symbolic and graphical representations. The post-test then evaluates students' ability to apply the resulting SPLDV knowledge.

Nevertheless, the findings should be generalized cautiously. The study was conducted at one school and focused on one Grade VIII topic. Student characteristics, teacher competence, technological facilities, and school culture may differ in other settings. Replication is necessary before concluding that similar results will occur across schools or mathematical topics.

The sample also consisted of intact classes rather than individually randomized students. This design is practical in school research but may allow pre-existing class differences to influence the results. The absence of reported pretest scores further limits causal interpretation.

The study measured learning outcomes primarily through an essay test administered after instruction. It did not measure motivation, anxiety, digital competence, collaboration, or long-term retention. Therefore, claims about these variables should be treated as preliminary observations unless supported by additional evidence.

Despite these limitations, the statistical results provide initial evidence that AI-based CTL supported by GeoGebra was associated with higher SPLDV learning outcomes than PjBL in the participating classes. The normality and homogeneity results supported the use of the t-test, while the reported t_{count} of 5.13 exceeded the critical value of 2.00.

Future research should employ a pretest–post-test design, report complete descriptive statistics, calculate effect sizes and confidence intervals, and include larger samples from multiple schools. Researchers should also examine implementation fidelity, student error patterns, learning motivation, and retention. Mixed-method approaches combining test scores, observation, interviews, and digital interaction records would provide a more comprehensive explanation of the intervention's effects.

Further studies may also compare GeoGebra-assisted CTL with conventional instruction, problem-based learning, discovery learning, or CTL without digital support. Such comparisons would help distinguish the contribution of contextual pedagogy from the contribution of dynamic visualization. A factorial design could examine whether CTL and GeoGebra produce independent or interactive effects.

In summary, the higher learning outcomes observed in the experimental class can plausibly be explained by the integration of meaningful contexts, active mathematical modeling, multiple representations, dynamic visualization, feedback, and teacher-guided reflection. These mechanisms provide a coherent pedagogical explanation for the statistical difference. However, stronger experimental controls and more complete statistical reporting are needed before broad causal conclusions can be established.

4. Conclusion

This study concluded that the first hypothesis was accepted, there was a significant difference in mathematics learning outcomes between students using AI-based CTL and PjBL (t_{count} 5.13 > t_{table} 2.00). The second hypothesis was also proven, the average learning outcomes of students with AI-based CTL were higher so it was stated to be more effective. Through GeoGebra, students more easily understand concepts, are skilled at solving problems, and are more motivated and active. However, the study was limited to one school, one class, and one material (SPLDV) so it cannot be generalized. For future research, it is recommended to expand the subjects, materials, and add other variables for more comprehensive results.

The findings provide empirical evidence that the integration of Contextual Teaching and Learning with GeoGebra-assisted mathematical visualization offers a promising instructional approach for teaching the Two-Variable Linear Equation System. Under the conditions of this study, students who participated in the AI-based CTL class obtained better mathematics learning outcomes than students who participated in the PjBL class. The statistical decision was supported by the calculated t-value of 5.13, which exceeded the critical value of 2.00. This result indicates that the difference between the two groups was unlikely to have occurred solely by chance according to the significance criterion applied in the study.

The higher learning outcomes obtained by the experimental class may be associated with the correspondence between the CTL model and the characteristics of SPLDV material. SPLDV requires students to identify unknown quantities, formulate relationships between variables, construct mathematical equations, perform algebraic procedures, and interpret the resulting solution. Contextual learning can support this process by

presenting mathematical relationships through situations that students can recognize and understand. As a result, equations are not introduced merely as abstract symbolic expressions but as mathematical representations of meaningful relationships.

The use of GeoGebra further strengthened this contextual learning process by providing dynamic connections among algebraic, numerical, and graphical representations. Students could observe that each linear equation was represented by a straight line and that the solution of the system corresponded to the intersection of the two lines. This visualization may have helped students connect the ordered pair obtained through elimination or substitution with its graphical meaning. It also provided an opportunity to verify calculations and identify inconsistencies between symbolic and graphical solutions.

The contribution of GeoGebra should nevertheless be understood as part of an integrated learning process. The improvement observed in the experimental class cannot be attributed exclusively to the application because the intervention also included contextual problems, teacher guidance, student discussion, mathematical modeling, questioning, and reflection. The instructional value of the technology emerged from how it was incorporated into the CTL stages. Therefore, the results support the combination of contextual pedagogy and dynamic mathematical visualization rather than the use of digital technology without an appropriate instructional framework.

The findings also emphasize the importance of multiple mathematical representations in algebra learning. Students need to understand SPLDV through contextual descriptions, symbolic equations, numerical values, and graphs. The ability to move among these representations can help students develop more flexible mathematical understanding. Students who understand only an algebraic procedure may be able to obtain a numerical answer but may not fully understand why the procedure works or what the answer means. The integrated learning model used in this study provided opportunities for students to connect these different forms of representation.

From a practical perspective, the AI-based CTL model can be considered an alternative for teachers who aim to improve the quality of SPLDV instruction. Teachers can begin the lesson with situations involving prices, quantities, transportation costs, or other relationships between two unknown values. Students can then identify the relevant information, define variables, formulate equations, and solve the system. GeoGebra can subsequently be used to visualize and verify the solution before students interpret it according to the original context. This sequence maintains the central role of mathematical reasoning while benefiting from digital visualization.

The findings do not imply that PjBL is generally ineffective or inferior in all mathematics-learning contexts. PjBL has important strengths in developing collaboration, creativity, project management, responsibility, communication, and the production of meaningful learning products. However, its effectiveness may depend on the instructional objectives, characteristics of the material, duration of learning, design of the project, and type of assessment. In the present study, the assessed outcome focused specifically on students' achievement in solving SPLDV essay questions. The advantage of AI-based CTL may therefore reflect its closer alignment with the contextual modeling, algebraic procedures, and graphical interpretation required by the assessment.

The study also demonstrates the importance of using an appropriate measurement instrument when comparing learning models. The instrument analysis indicated that the test items met the reported validity criteria and obtained a reliability coefficient of 0.83. The predominance of questions with moderate difficulty provided students with reasonable opportunities to demonstrate differences in their SPLDV competence. These characteristics support the use of the instrument for assessing learning outcomes, although future studies should clarify the inconsistency between the reported five essay questions and the six items presented in the validity and difficulty analyses.

The fulfillment of the normality and homogeneity assumptions also supported the use of the parametric statistical procedure. Nevertheless, the statistical evidence presented in this study should be strengthened by more complete reporting. Future studies should include the mean, standard deviation, sample size, degree of freedom, exact probability value, confidence interval, and effect size. These statistics would enable readers to evaluate not only whether the difference was statistically significant but also whether it was educationally substantial.

The conclusion regarding effectiveness should also be interpreted within the boundaries of the research design. The study compared intact classes selected through cluster random sampling, but it did not report pretest results demonstrating that the experimental and control groups had equivalent SPLDV abilities before instruction. Consequently, the higher post-test performance of the experimental class may be associated with the intervention, but pre-existing differences between classes cannot be ruled out completely. A pretest–post-test control-group design would provide stronger evidence regarding changes attributable to the learning model.

Another limitation concerns the scope of the

participants and research setting. The study was conducted at SMP Negeri 4 Gadingrejo with Grade VIII students and focused only on SPLDV material. The participating students shared a particular school environment, curriculum implementation, level of technological access, and learning culture. Therefore, the results should not be generalized automatically to students from different schools, regions, grade levels, or mathematical topics. Replication in more diverse educational settings is needed to evaluate the consistency of the findings.

CRediT authorship contribution statement

Yopi Ayomi Afandi: Conceptualization, Methodology, Software, Investigation, Formal analysis, Writing - original draft. **Noerhasmalina:** Supervision, Validation, Writing - review & editing. **Ana Istiana:** Supervision, Methodology, Validation. **Lisdwiana Kurniati:** Resources, Data curation, Visualization. **Korhan Cengiz:** Formal analysis, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgment

The authors gratefully acknowledge SMP Negeri 4 Gadingrejo, Lampung, for supporting the implementation of this research. The researcher would also like to thank all parties involved for their assistance and cooperation from the beginning of the research to its completion.

References

1. Asyhara, S. A., Yuniarti, T., and Asmiati, A. (2022). Pengembangan Bahan Ajar Dengan Metode Pq4R Untuk Menanamkan Kemampuan Pemahaman Konsep Matematis, *Hipotenusa Journal of Research Mathematics Education (HJRME)*, Vol. 1, No. 1, 1–13. doi:10.36269/hjrme.v1i1.803.
2. Efendi, D., and Caswita, C. (2018). Berpikir Kritis Matematis Dengan Model Pembelajaran Penemuan, *Hipotenusa: Journal of Research in Mathematics Education*, Vol. 1, No. 1, 30–44. doi:10.36269/hjrme.v1i1.817.
3. Sari, V. A., Stiyaningsih, A., and Subehi, M. (2020). Efektivitas Model Pembelajaran Kooperatif Tipe STAD Terhadap Hasil Belajar Matematika Siswa Kelas X MIPA 1 SMA Negeri 1 Kajen, *Penelitian Pendidikan Matematika*, Vol. 7, No. 3, 127–140.
4. Kemendikbudristek. (2023). *PISA 2022 Dan Pemulihan Pembelajaran Di Indonesia*, Kemendikbudristek.
5. Angriani, A. D., Kahar, A. K., Aini, N., Prastika, F., Zaky, A., Arafah, Y. Y., Tenri, A., and Sompa, M. (2025). Development of Interactive Learning Media Assisted By Smart Apps Creator 3 in Spldv Materials To Increase Students' Learning Interest, *Hipotenusa Journal of Research Mathematics Education*, Vol. 8, No. 2, 2723–486.
6. Muslihah, N. N., and Suryaningrat, E. F. (2021). Model Pembelajaran Contextual Teaching and Learning Terhadap Kemampuan Pemecahan Masalah Matematis, *Plusminus: Jurnal Pendidikan Matematika*, Vol. 1, No. 3, 553–564. doi:10.31980/plusminus.v1i3.1445.
7. Mirdad, J., and Pd, M. I. (2020). Model-Model Pembelajaran (Empat Rumpun Model Pembelajaran), Vol. 2, No. 1, 14–23.
8. Dhani, M. I., and Rahayu, W. (2023). Literatur Review: Contextual Teaching and Learning (CTL) Dalam Pembelajaran Matematika, *Jurnal Ilmiah IKIP Mataram*, Vol. 10, No. 2, 118–135.
9. Hasan, H. (2021). Meningkatkan Hasil Belajar Matematika Melalui Penerapan Model Contextual Teaching and Learning Pada Era New Normal, *Indonesian Journal of Educational Development*, Vol. 1, No. 4, 630–640. doi:10.5281/zenodo.4560726.
10. Yasin, M., Huda, S., Putra, F. G., Syazali, M., Umam, R., and Widyawati, S. (2020). IMPROVE Learning Model and Learning Independence : Influence and Interaction on Mathematics Problem-Solving Abilities in Islamic Boarding School IMPROVE Learning Model and Learning Independence : Influence and Interaction on Mathematics Problem-Solving Ab, *Journal of Physics: Conference Series*, Vol. 1467, 012003. doi:10.1088/1742-6596/1467/1/012003.
11. Asngari, D. R. (2015). Penggunaan Geogebra Dalam Pembelajaran Geometri, *Seminar Nasional Matematika Dan Pendidikan Matematika UNY*, 299–302.
12. Sugiyono. (2019). *Metode Penelitian Kuantitatif, Kualitatif, Dan R\&D*, Alfabeta.
13. Putra, F. G. et al. (2024). Mathematical Representation: A Bibliometric Mapping of the Research Literature (2013--2022), *Infinity Journal of Mathematics Education*, Vol. 13, No. 1, 1–26. doi:10.22460/infinity.v13i1.p1-26.
14. Huda, S., Anggraini, L., Saputri, R., Syazali, M., and Umam, R. (2019). Learning Model to Improve the Ability to Understand Mathematical Concepts, *Prisma*, Vol. 173, No. 2, 173–181.
15. Lestari, K. E., and Yudhanegara, M. R. (2015). *Penelitian Pendidikan Matematika*, Refika Aditama.

16. Huda, S., Suherman, Komarudin, Syazali, M., and Umam, R. (2020). The Effectiveness of Al-Qurun Teaching Model (ATM) Viewed from Gender Differences, *Journal of Physics: Conference Series* (Vol. 1467), 12001. doi:10.1088/1742-6596/1467/1/012001.

17. Wang, Y. F., Zhang, Y. J., Fu, L. H., and Hongwei, L. (2021). Three-Dimensional Gravity Inversion Based on 3D U-Net++, *Applied Geophysics*, Vol. 18, No. 4, 451–460. doi:10.1007/s11770-021-0909-z.

18. Sumarni, S., Ramadhani, R., Sazaki, Y., Tria, R., Andika, W. D., and Prasetiyo, A. E. (2019). Development of Child Friendly ICT Textbooks to Improve Professional Competence of Teacher Candidates, *Journal of Education and Gifted Young Scientists*, Vol. 7, No. 3, 643–658. doi:10.17478/jegys.596095.

19. Prastowo, R., Huda, S., Umam, R., Jermsittiparsert, K., and Eko, A. (2019). Academic Achievement and Conceptual Understanding of Electrodynamics, *Jurnal Ilmiah Pendidikan Fisika Al-Biruni*, Vol. 8, No. 2, 165–175. doi:10.24042/jipfalbiruni.v0i0.4614.

20. Sumarni, S. et al. (2019). Behavior in Early Childhood (2--3) Years: A Case Study on the Use of Gadgets, *International Journal of Innovation, Creativity and Change*, Vol. 8, No. 8, 384–404.

21. Munifah, M., Romadhona, A. N., Ridhona, I., Ramadhani, R., Umam, R., and Tortop, H. S. (2019). How to Manage Numerical Abilities in Algebra Material?, *Al-Jabar: Jurnal Pendidikan Matematika*, Vol. 10, No. 2, 223–232.

22. Riduwan. (2015). *Metode Dan Teknik Menyusun Proposal Penelitian*, Alfabeta.

23. Nasution, S. H. (2018). Penggunaan GeoGebra Untuk Membuat Soal Interaktif Matematika, *MSOpen*, 1–13.

24. Rahmawati, T. D., Wahyuningsih, W., and Dua Getan, M. A. (2019). Pengaruh Model Pembelajaran Contextual Teaching and Learning Terhadap Hasil Belajar Matematika, *JINoP: Jurnal Inovasi Pembelajaran*, Vol. 5, No. 1, 83. doi:10.22219/jinop.v5i1.8021.

25. El Baba, M., Kayastha, P., Huysmans, M., and De Smedt, F. (2020). Evaluation of the Groundwater Quality Using the Water Quality Index and Geostatistical Analysis, *Water*, Vol. 12, No. 262, 1–14. doi:10.3390/w12010262.

26. Ridwanulloh, M. U., Huda, S., and Umam, R. (2022). Innovative Leadership Management: The Pattern of School Quality Development, *Progressiva: Jurnal Pemikiran Dan Pendidikan Islam*, Vol. 11, No. 1, 25–42. doi:10.22219/progresiva.v11i01.20742.

27. Saregar, A. et al. (2022). Natural Disaster Education in School: A Bibliometric Analysis, *International Journal of Educational Methodology*, Vol. 8, No. 4, 743–757. doi:10.12973/ijem.8.4.743.