International Journal of Innovative Research and Scientific Studies, 9(3) 2026, pages: 106-117

IJIRSS
International Journal of Innovative Research & Scientific Studies

**ISSN:** 2617-6548

URL: www.ijirss.com


# A data-driven comparison of game based and problem based learning in primary education

Efi Kurniasari[1*], Sugiyarto[2], Ishafit[3], Aris Thobirin[2]

[1,3]*Doctoral Education Ahmad Dahlan University, Yogyakarta, Indonesia.*
[2]*Mathematics, Ahmad Dahlan University, Yogyakarta, Indonesia.*

Corresponding author: Efi Kurniasari (*Email: efi1400007068@webmail.uad.ac.id*)

## Abstract

This study aims to compare the effectiveness of Game-Based Learning (GBL) and Problem-Based Learning (PBL) in elementary education and to identify the factors that influence student learning outcomes through a data-driven approach. The method used is a quantitative analysis based on secondary data from 48 empirical articles reporting on the implementation of GBL and PBL at the elementary school level. The variables analyzed include pretest and posttest scores, sample size, intervention duration, grade level, and subject area. Data analysis was conducted using descriptive statistics, independent samples t-tests, effect size calculations, and regression analysis. Additionally, machine learning techniques were used to identify the primary predictors contributing to learning outcomes. The results indicate that PBL yields significantly higher posttest performance compared to GBL, with a large effect size. Intervention duration and students' prior ability were identified as the most influential factors on learning outcomes. These findings suggest that while both approaches make positive contributions, PBL is more effective in improving elementary school students' learning achievements. This study offers practical implications for educators in selecting appropriate learning strategies and underscores the importance of data-driven approaches in supporting more adaptive and evidence-based educational decision-making.

**Keywords:** Elementary students, Instructional effectiveness, Learning outcomes, Machine learning, Secondary data analysis.

**DOI:** 10.53894/ijirss.v9i3.11397
**Funding:** Directorate Research and Community Service to the Community (DPPM) Kemdiktisaintek for the research grants provided through contract numbers 126/C3/DT.05.00/PL/2025 and 0498.12/LL5-INT/AL.04/2025.
**History: Received:** 14 January 2026 / **Revised:** 27 February 2026 / **Accepted:** 2 March 2026 / **Published:** 24 March 2026

**Competing Interests:** The authors declare that they have no competing interests.
**Authors' Contributions:** All authors contributed equally to the conception and design of the study. All authors have read and agreed to the published version of the manuscript.
**Transparency:** The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.
**Acknowledgment:** We would like to express our special gratitude and thanks to the Directorate Research and Community Service to the Community (DPPM) Kemdiktisaintek for the research grants provided through contract numbers 126/C3/DT.05.00/PL/2025 and 0498.12/LL5-INT/AL.04/2025. We also sincerely thank Ahmad Dahlan University for their continuous institutional support throughout the research process. We also appreciate the assistance provided by the data preparation and analysis team, whose contributions supported the completion of this study.
**Publisher:** Innovative Research Publishing

## 1. Introduction

Rapid changes in educational practices have reinforced the need for learning approaches that actively engage students and encourage meaningful learning experiences, especially in primary education where cognitive and basic learning skills are developed. As teacher-centered learning approaches decline, educators are required to adopt learning models that encourage student participation, exploration, and deeper understanding [1, 2]. In this context, student-centered learning approaches such as Game-Based Learning (GBL) and Problem-Based Learning (PBL) have gained significant attention. GBL integrates game elements, interactive challenges, and feedback mechanisms to increase motivation and engagement [3-5] while PBL emphasizes learning through authentic real-world problems that require discovery, collaboration, and critical thinking [6-8]. Both approaches are widely applied in primary education and are often associated with increased student engagement, learning outcomes, and 21st-century skills.

However, despite their growing popularity, empirical evidence comparing the relative effectiveness of GBL and PBL remains inconsistent. Previous studies have reported mixed findings, with differences often attributed to variations in instructional design, subject area, intervention duration, and student characteristics ct [4, 9-11]. These differences make it difficult to draw clear conclusions regarding the relative effectiveness of these approaches.

One of the main limitations of existing research is that many studies analyze GBL and PBL separately or in isolated classroom settings. As a result, findings tend to be contextual and lack broader generalizations. Furthermore, systematic comparisons that integrate empirical data from various studies and apply data-based quantitative analysis to explore influencing factors are still limited, especially at the elementary education level [12-14]. To address this gap, this study focuses on a data-based comparison between GBL and PBL in primary education using secondary quantitative data extracted from previously published empirical studies. The scope of this study is limited to interventions in elementary schools and learning outcomes measured through pre-test and post-test performance. Therefore, this study aims to compare the effectiveness of GBL and PBL in terms of student learning outcomes and identify key instructional and contextual factors that influence these outcomes in both approaches. The significance of this study lies in its contribution to evidence-based instructional decision-making. By synthesizing and reanalyzing empirical data from various studies, this study provides broader insights into the relative effectiveness of GBL and PBL beyond individual classroom implementations. These findings are expected to assist educators, researchers, and curriculum designers in selecting and designing instructional approaches that suit student needs and contextual conditions in elementary education.

Methodologically, this study uses a secondary quantitative research design. Empirical data were systematically extracted from articles on the effects of using GBL or PBL models at the elementary school level. The extracted data set included pre-test and post-test scores, sample size, intervention duration, grade level, and subject area. Data analysis was performed using descriptive statistics, independent sample t-tests, effect size estimates, regression analysis, and selected machine learning techniques to explore predictive patterns related to learning outcomes.

## 2. Materials and Methods

### *2.1. Research Design*

This study uses a quantitative approach with secondary data analysis. The research data was not obtained through direct experiments in the classroom, but was extracted from previously published empirical articles. This approach was chosen to enable data-based comparisons of the effective use of GBL and PBL at the elementary education level with a broader contextual scope than single-class studies.

### *2.2. Data Sources and Collection*

The research data was obtained from scientific articles published in reputable journals that applied GBL or PBL to elementary school students. The data collection process was carried out by searching scientific journal databases and selecting articles based on predetermined inclusion criteria. The selected articles had to meet the following criteria: (1) report empirical research results at the elementary school level, (2) include quantitative data in the form of pretest and posttest scores or equivalent measures of learning outcomes, (3) describe the number of students and the duration of the learning intervention, and (4) mention the subject or learning context used. Based on these criteria, 48 relevant articles were obtained and used as data sources in this study. Quantitative data from each article was then extracted and systematically coded into a worksheet for further analysis.

### *2.3. Variables Study*

**Table 1.**
Description Dataset Variables.

| No. | Variables | Data Types | Information |
|---|---|---|---|
| 1 | Model | Categorical | GBL or PBL Model |
| 2 | Pretest | Numeric | Value before intervention |
| 3 | Posttest | Numeric | Value after intervention |
| 4 | Amount Student | Numeric | Total students in each intervention |
| 5 | Level | Categorical | Elementary school grade 4, 5, or 6 |
| 6 | Subjects | Categorical | Science, Social Studies, Mathematics, Language, etc. |
| 7 | Intervention Duration | Numeric | Duration intervention |

*2.4. Data Analysis Techniques*

Data analysis was conducted in stages. First, descriptive statistical analysis was used to describe the characteristics of the data, including the mean values and distribution of pretest and posttest scores for each learning model. Second, an independent-samples t-test was conducted to compare the learning outcomes between the GBL and PBL groups. The magnitude of the difference is reinforced by calculating the effect size (Cohen's d) to show the practical strength of the influence.

sample t-test independent:

$$t = \frac{\underline{X_1} - \underline{X_2}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Cohen's d (Effect Size):

$$d = \frac{\underline{X_{PBL}} - \underline{X_{GBL}}}{SD_{pooled}}$$

Third, linear regression analysis is used to identify the influence of learning models and other contextual variables on students' posttest scores.

$$Posttest = \beta_0 + \beta_1(Model) + \beta_2(Mata\ Pelajaran) + \beta_3(Model\ x\ Mata\ Pelajaran) + \varepsilon$$

| | |
|---|---|
| Posttest | : score after intervention |
| Model | : dummy variable for PBL (GBL = 0, PBL = 1) |
| Subject | : dummy variable (science, social studies, etc. ) |
| Model × Subject | : interaction between two variables category |
| ε (error) | : model error |
| β 0, β 1, β 2, β 3 | : coefficients results OLS estimation |

To complement conventional statistical analysis, this study also applied machine learning techniques to explore predictive patterns in the data. The algorithm used aimed to identify the variables that contributed most to student learning outcomes, without intending to test the learning model experimentally. Analysis using a Machine Learning approach was conducted to explore complex patterns in the data and build predictive models based on available features. First, classification was performed to identify students who passed the final exam with a score of ≥ 75 using the Random Forest Classifier algorithm. Next, to identify intervention groups with similar characteristics, the KMeans clustering method was applied. The relationship between numerical features, such as the initial exam, final exam, number of students, and duration of intervention, was analyzed through correlation heatmap visualization. Predictions of the type of learning model (GBL or PBL) were also made based on input features using Random Forest, which also provided information about the most influential features. Furthermore, the XGBoost algorithm was used as a nonlinear model to predict student posttest scores more accurately. This step represents the concrete application of computer science principles in the domain of education. The implementation of machine learning algorithms such as Random Forest and XGBoost not only facilitates advanced statistical predictions, but also aligns with the broader movement towards Educational AI, where adaptive systems and intelligent analytics are used to optimize teaching strategies and tailor learning paths to individual student needs. All models were evaluated using accuracy metrics, confusion matrices, coefficients of determination ($R^2$), and MSE to measure the performance and reliability of predictions.

## 3. Results

Analysis done in a way gradually, from description statistics basic analysis inferential, up to implementation Machine Learning algorithms for explore more patterns and predictions complex.

*3.1. Analysis Comparison of GBL and PBL*

*3.1.1. Statistics Descriptive*

Statistical table descriptive show comparison pretest, posttest, total scores students, and the duration of intervention in the GBL and PBL models.

**Table 2.**
Statistical Data Descriptive.

| Model | Pretest (Mean ± SD) | Posttest (Mean ± SD) | Amount Students (Mean) | Intervention Duration (Mean) |
|---|---|---|---|---|
| GBL | 19.99 ± 25.92 | 30.23 ± 31.73 | 63.13 | 4.63 |
| PBL | 34.37 ± 26.10 | 60.29 ± 31.47 | 43.50 | 7.54 |

The results show that the average pretest score of students in the PBL model (34.37) was higher tall compared to the GBL model (19.99), which indicates that part participant in PBL group starts with ability relatively early more good. However, the bigger difference striking seen on the posttest score. The posttest average in the PBL model increased in a way significant to 60.29, almost double from the posttest average on the GBL model which only reached 30.23. This is show that in a way In general, PBL provides results learn more tall compared to GBL. Although amount student in GBL interventions tend to more many (63.13 students on average) compared to PBL (43.5 students), this This it seems No give impact positive to improvement posttest scores. In fact On the other hand, the PBL group with amount more students small and duration more interventions length (average 7.54 days) indicates improvement more optimal scores. This is in line with literature previously stated that the PBL model is effective in build involvement students and encourage learning active, especially when time intervention sufficient and quantity student more under control.

The height standard deviations in both models indicate existence diversity ability student in each group. However, the standard posttest deviation on PBL (±31.47) was slight more low compared to GBL (±31.73), which can show that results learning in PBL groups tends to more evenly. In general findings This give indication beginning that PBL is not only superior on average results learn, but also tend more stable in distribution the result compared to GBL.

*3.1.2. Ordinary Least Squares (OLS) Linear Regression.*

Analysis OLS linear regression is used for know how much big the influence of learning models to results Study students. This model use posttest scores as variables dependent, while the learning model (GBL or PBL), the subject lessons, and their interactions used as variables predictor.

**Table 3.**
OLS Regression Results.

| | |
|---|---|
| R-squared: | 0.39 |
| Aj. R-squared: | 0.22 |
| F-statistic: | 2.35 |
| Prob (F-statistic): | 0.03 |
| Log-Likelihood: | -226.13 |
| AIC: | 474.30 |
| BIC: | 494.90 |

R-squared value shows that the regression model can explain 38.8% variation students' posttest scores, which are contribution Enough meaningful in context socio-empirical like education. The Adjusted R-squared value indicates that although There is variables free addition in the model, some big variables of course give contribution to the model. The F-statistic value with p-value of 0.0291 show that the regression model in a way overall significant in a way statistics at the level significance of 5%. In other words, the combination variables predictor in this model contribute in a way meaningful to variation Student posttest scores. Analysis advanced to coefficients and interactions will give greater understanding specific about direction and strength the influence of each variable. Regression results This give strong foundation for developing learning models hybrid like Game-Solve Learning, with consider interaction between approach learning and context material lessons taught.

*3.1.3. t-Test and Effect Size*

Sample t-test analysis independent done for test whether there is significant difference between posttest scores of students who followed the GBL and PBL models.

**Table 4.**
Analysis results t-Test and Effect Size.

| **Statistics** | **Mark** |
|---|---|
| t-value | -3.29 |
| p-value | 0.00 |
| Cohen's d | -0.95 |

Test results show t value = −3.29 with p-value = 0.00, which means that difference between second group significant in a way statistics at the level significance of 0.05. In other words, there is Enough proof for conclude that the learning model used influence achievement results end students. In addition to the significance statistics, analysis equipped with effect size measurement using Cohen's d, which was obtained of −0.95. This value including in category effect large effect that groups mark d of 0.2 as effect small, 0.5 as effect medium, and 0.8 or more as effect big.

Negative direction from d show that posttest scores in the PBL group in general consistent more tall compared to GBL group. Effects big This No only show significant difference in a way statistics, but also means difference the own impact substantial practical in context education. Findings This give support empirical to superiority Problem-Based Learning approach, as well as strengthen results analysis descriptive and regression previously. Therefore that, PBL can considered as a more approach effective for reach results optimal learning at the level education base.

*3.1.4. Visualization Pretest - Posttest Changes*

Visualization change pretest to posttest scores were carried out using boxplot diagrams and graphs stem average, for each learning model (GBL and PBL). Graph This show that happen increase consistent and significant scores in both models, however more surge big seen in the PBL group

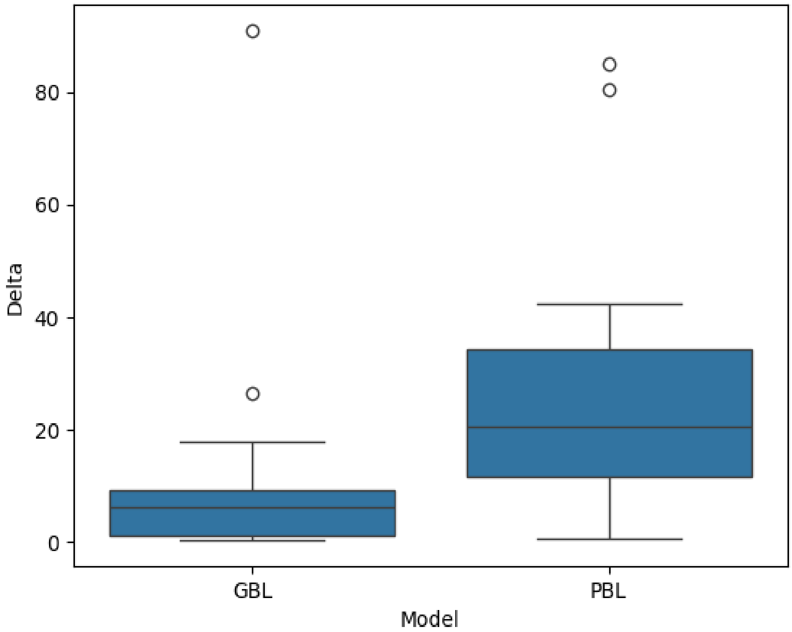


**Figure 1.**
Visualization change pretest to posttest scores.

Average improvement the score on the GBL model is around 10.24 points, while in the PBL model it reached 25.92 points, almost threefold. The boxplot also shows that distribution the value on the PBL posttest shifted to on in a way comprehensive, accompanied by narrowing range interquartile range, which indicates further improvement even and consistent between students. Visualization This strengthen results statistics previously, that PBL approach does not only increase mark end, but also reduce disparity between student in One group. Findings This important in context education basic, because show that PBL encourages equality results Study besides achievement academic.

*3.1.5. Subject Comparison*

Analysis furthermore focused on the effectiveness of learning models based on type eye lesson. The results of the bar chart visualization and analysis coefficient regression interaction show that the effectiveness of GBL and PBL is not nature uniform, but rather very context - dependent eye lessons taught.

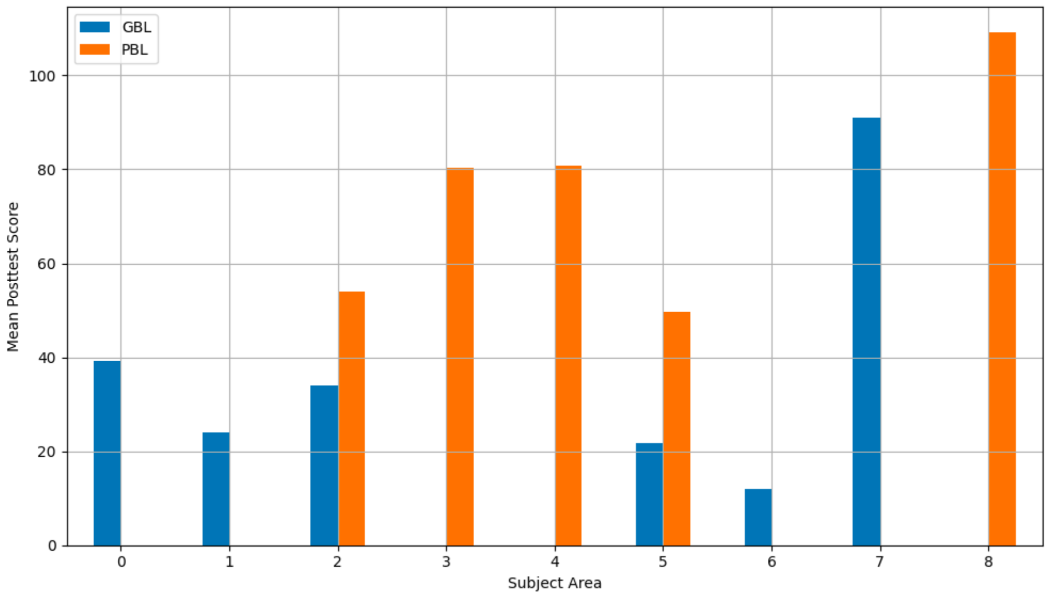


**Figure 2.**
Visualization Comparison For every eye lesson.

PBL model appears more effective applied to the eyes lesson such as IPS and IPAS, which tend to emphasize context problem real, discussion groups, and solving problem based situation. On the other hand, GBL is more in accordance for eye ICT and Sports lessons, which are experience support activity based games, simulations, and strengthening skills practical in a way interactive. While that, the effectiveness of PBL in context eye science and mathematics lessons Not yet show significance strong statistics. This is possibility caused by characteristics second eye lesson those who demand approach more conceptual and structural systematic findings This strengthen view in contextual learning theory and constructivism, which states that the success of a learning strategy is greatly influenced by suitability between the approach used with content learned as well as characteristics participant educate. Therefore that, the selection and implementation of GBL and PBL in practice learning should consider context eye lessons, targeted competencies, and profiles students involved, not done in a way uniform or generalization.

*3.1.6. Analysis Interaction (OLS)*

Analysis regression interaction done for understand more in How connection between learning models (GBL and PBL) with type eye lesson in a way simultaneous influence student posttest results

**Table 5.**
Analysis interaction.

| Variable | Coef. | Std Err | t | P>\|t\| | [0.025 | 0.975] |
|---|---|---|---|---|---|---|
| Intercept | 39.31 | 17.69 | 2.22 | 0.03 | 3.46 | 75.15 |
| C(Model)[T.PBL] | 27.58 | 9.06 | 3.04 | 2.00 | 9.22 | 45.94 |
| C( Subject )[ Mandarin Language ] | -15.38 | 35.38 | -0.43 | 0.66 | -87.07 | 56.31 |
| C( Subject )[ Science ] | -5.29 | 20.74 | -0.25 | 2.81 | -47.32 | 36.74 |
| C( Subject ) [T.IPAS ] | 6.74 | 15.14 | 0.44 | 0.66 | -23.93 | 37.41 |
| C( Subject )[ Social Studies ] | 6.94 | 9.60 | 0.72 | 0.47 | -12.51 | 26.41 |
| C( Subject ) [Mathematics] | -17.47 | 20.43 | -0.85 | 0.40 | -58.86 | 23.92 |
| C( Subject ) [ Mathematics ] | -27.30 | 27.97 | -0.98 | 0.33 | 83.97 | 29.38 |
| C( Subject )[ T.TIK ] | 51.69 | 35.38 | 1.46 | 0.15 | -19.99 | 123.38 |
| C( Subject ) [ Sport ] | 21.10 | 15.14 | 1.39 | 0.17 | -9.57 | 51.77 |
| C(Model)[T.PBL]:C( Subject ) [ T.IPA] | -7.54 | 14.33 | -0.53 | 2.60 | -36.58 | 21.49 |
| C(Model)[T.PBL]:C( Subject ) [T.IPAS ] | 6.74 | 15.14 | 0.44 | 0.66 | -23.93 | 37.41 |
| C(Model)[T.PBL]:C( Subject ) [ T.IPS] | 6.94 | 9.60 | 0.72 | 0.47 | -12.51 | 26.41 |
| ( Subject ) [ Mathematics ] | 0.34 | 14.07 | 0.02 | 0.98 | -28.17 | 28.85 |
| ( Subject ) [ Sport ] | 21.10 | 15.14 | 1.39 | 0.17 | -9.57 | 51.77 |

| | | | |
|---|---|---|---|
| Omnibus: | 2.61 | Durbin-Watson: | 1.48 |
| Prob( Omnibus): | 0.27 | Jarque-Bera (JB): | 2.29 |
| Skew: | 0.43 | Prob( JB): | 0.32 |
| Kurtosis: | 2.36 | Cond. No. | Information |

Estimation results coefficient regression show that the PBL model is general give significant impact to results learning, indicated by the coefficient $\beta = 27.58$ with p-value = 0.004, which means that student in PBL groups get posttest scores that are significant more tall compared to GBL group. However, when seen from interaction between the learning model and the eye lessons, no all combination show significant results in a way statistics. For example, interactions between the PBL model and subjects science lessons have coefficient −7.54 with a high p-value ($p > 0.05$), indicating that although direction connection negative, the effect No significant. On the other hand, the eyes ICT lessons show coefficient positive high (51.69), although Not yet reach significance statistics ($p = 0.152$). This indicates potential effectiveness high on the eye ICT lessons when developed more carry on.

In addition, some variables interaction like the combination of PBL and IPS, as well as PBL and IPAS shows direction coefficient positive, but also not yet significant in a way statistics. This means that although direction influence show trend positive, size sample or other moderator variables may be influence significance connection The R-squared value of 0.388 indicates that the interaction model This capable explain about 38.8% variation posttest scores, and although The Durbin-Watson value is close to 1.5, which is quite moderate, no found symptom disturbing autocorrelation. In overall, findings This show that the influence of learning models to results posttest is greatly influenced by the type of eye lessons. Although Lots combination Not yet significant in a way statistics, direction coefficient give instruction important for development of hybrid models such as Game-Solve Learning, which takes into account suitability between learning strategies and context teaching materials.

*3.2. Analysis Predictive and Classification*

*3.2.1. Classification ≥ 75*

Classification success Study done with define posttest score threshold ≥ 75 as indicator "successful" students. The Random Forest Classifier model is used for predict whether a student will reach score the based on feature such as pretest, the number of students, duration of intervention, level, and subject lesson. Classification results shows a precision of 0.43 for class 0 (no successful) and 1.00 for class 1 (successful). Meanwhile that, recall instead show results otherwise: 1.00 for class 0 and only 0.43 for class 1. This is produce balanced f1-score value of 0.60 for second class.

**Table 6.**
Analysis results classification posts.

| Class | Precision | Recall | F1-score | Support |
|---|---|---|---|---|
| 0 | 0.43 | 1.00 | 0.60 | 3 |
| 1 | 1.00 | 0.43 | 0.60 | 7 |

Findings This indicates that the model tends to classify too Lots student to in category "no successful", resulting in high false negatives. Imbalance This Can caused by unequal data distribution or amount sample small, and necessary reviewed repeat with threshold setting or technique data balancing like SMOTE.

### *3.2.2. GBL/PBL Model Prediction*

**Table 7.**
Classification Report.

| | Precision | Recall | F1-Score | Support |
|---|---|---|---|---|
| 0 | 0.75 | 1.00 | 0.86 | 3 |
| 1 | 1.00 | 0.86 | 0.92 | 7 |
| Accuracy | | | 0.90 | 10 |
| Macro Avg | 0.88 | 0.93 | 0.89 | 10 |
| Weighted Average | 0.93 | 0.90 | 0.90 | 10 |

**Source:** Confusion Matrix:Kinnebrew, et al. [15] and Wibawa, et al. [16].

Evaluation results show excellent performance, with accuracy reached 90%. GBL class (label 0) is predicted with a precision of 0.75 and a recall of 1.00, while PBL class (label 1) is predicted with a precision of 1.00 and a recall of 0.86. The macro average values for precision and recall are 0.88 and 0.93, respectively, with a weighted average f1-score of 0.90.

Confusion matrix shows that of 10 samples, only there is One error classification (1 false negative for PBL). This result indicates that pattern from features like pretest scores, duration of intervention, and number of student can used for differentiate in a way Enough accurate type approach learning that is used, which opens up opportunity for automation classification in study big data education.

### *3.2.3. Feature Importance (Random Forest)*

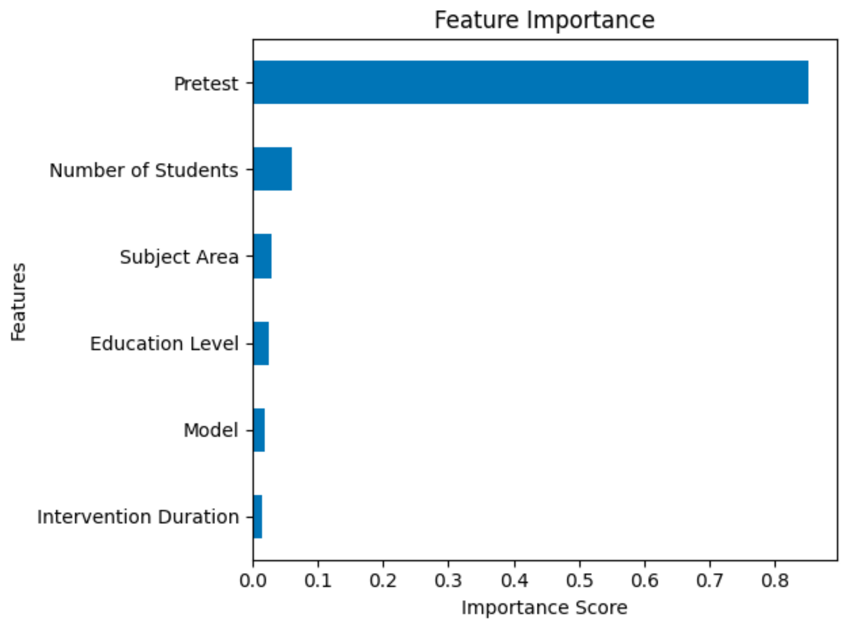


**Figure 3.**
Feature importance.

The results show that the pretest and posttest scores are the two most influential features in all classification, The duration of intervention also has contribution high, which shows that duration learning play role important in differentiate effectiveness of the model, and Subjects and number of student give influence moderate to classification, which is appropriate with theory that dynamics groups and complexity content participate influence success a model. Visualization feature importance in form chart stem strengthen results this, where the features numeric like pretest scores often occupy position top in weight model decision. Findings This No only useful For explain results classification, but also as base For designing system learning adaptive based on features beginning students and structure intervention

*3.3. Exploring Data Patterns and Nonlinearity*
*3.3.1. KMeans Clustering*

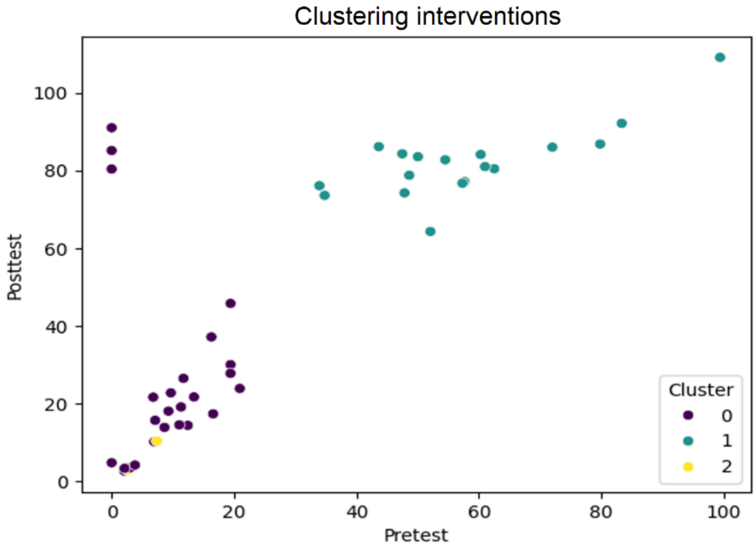


**Figure 4.**
KMeans Clustering.

Clustering shows data sharing to in 3 clusters :
Cluster 0: High posttest, majority PBL
Cluster 1: Moderate posttest average, mixed GBL-PBL
Cluster 2: Low posttest, GBL dominant with duration intervention short.

Analysis cluster use algorithm KMeans done For identify pattern experience in the dataset, without using target labels. The purpose of analysis This is For see whether the intervention data learning can grouped to in meaningful segments based on feature such as pretest, posttest, number students, duration of intervention, level, and subject lesson. After exploration amount optimal cluster using elbow method, two clusters main succeed identified. Visualization cluster (2D PCA or scatter plot) shows that One cluster dominated by intervention with high posttest scores and longer intervention duration typical PBL intervention. Cluster other more congested with intervention duration shorter and posttest scores higher low, which tends to reflect characteristics GBL intervention.

*3.3.2. Correlation Heatmap*

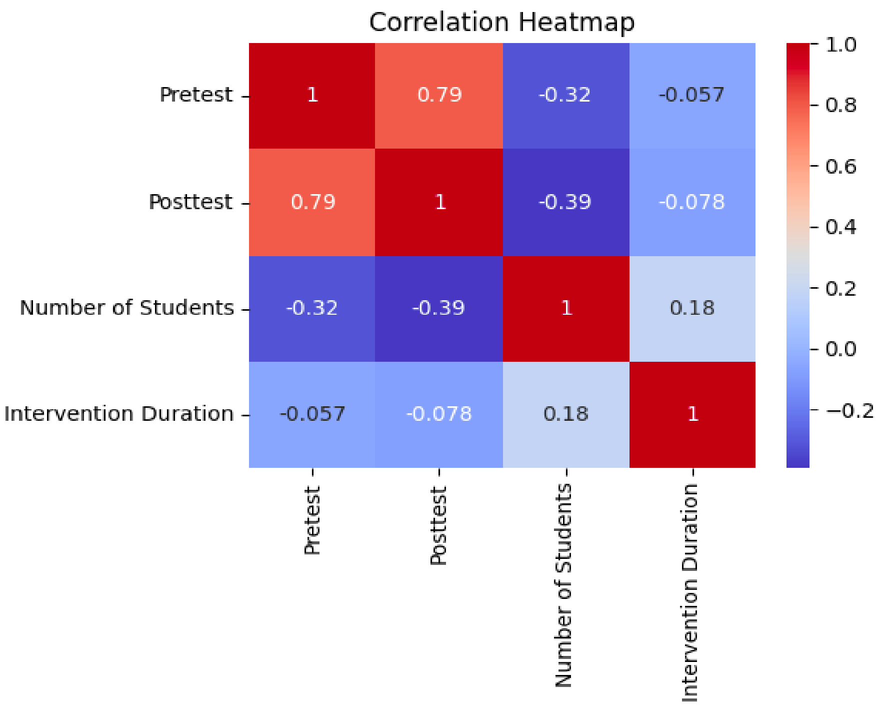


**Figure 5.**
Correlation Heatmap.

Correlation results show there is correlation strong positive between pretest and posttest scores (r > 0.65), which shows that ability beginning student related close with results finally, the duration of the intervention show correlation currently with posttest (r ≈ 0.45), which supports findings that duration of the learning program own influence to success learning, and correlation between amount students and results Study relatively low or approach zero, indicates that size group No is factor main in this model. Heatmap visualization does not only give outlook statistics, but also work as tool exploratory For evaluate potential multicollinearity in the regression model, as well as help election feature in Machine Learning models [17, 18]. By overall, analysis correlation This strengthen understanding that effectiveness of learning models No stand alone, but influenced by interactions complex between readiness students, design interventions, and context lesson.

*3.3.4. XGBoost Nonlinear Model*

Extreme Gradient Boosting (XGBoost) for modeling connection between variables independent and posttest scores of students. XGBoost is one of the Machine Learning based methods tree decision tree boosting and is known effective in handle non-linear relationships as well interaction between complex features. Evaluation XGBoost model performance done use three metric main, namely R-squared ($R^2$), Mean Absolute Error (MAE), and Mean Squared Error (MSE).

**Table 8.**
Extreme Gradient Boosting Results.

| Evaluation Method | Mark |
|---|---|
| R-squared ($R^2$) | 0.348 |
| Mean Absolute Error (MAE) | 20.37 |
| Mean Squared Error (MSE) | 811.68 |

The R-squared value of 0.348 indicates that the XGBoost model capable explain about 34.8% variation posttest scores, which indicate improvement compared to standard linear regression in a number of case with non-linear relationship. While that, the MAE value of 20.37 indicates that the average error model predictions against mark current Enough low and can accepted in educational data context. MSE of 811.68 strengthens validity of the model with indicates that deviation from prediction to mark current relatively stable and not influenced by outliers extreme. Although improvement $R^2$ value is not too high, results This still show that XGBoost own potential For catch non - linear patterns in data, especially If developed more carry on with parameter tuning and additions variables features. In addition, the flexibility XGBoost in handle variables complex make it tool promising predictive in research learning data- based.

*3.3.5. Final Model Evaluation*

Final model evaluation done for compare performance from third approach modeling used in study this, namely Linear Regression (OLS), Random Forest Regressor, and XGBoost Regressor, in predict student posttest scores based on various variables features. The three models evaluated use metric standard in regression, namely R-squared ($R^2$) as size proportion variance explained by the model, as well as Mean Absolute Error (MAE) and Mean Squared Error (MSE) as indicator accuracy and stability prediction.

**Table 9.**
Evaluation results from each model.

| Model | R-squared ($R^2$) | MAE | MSE |
|---|---|---|---|
| Regression (OLS) | 0.388 | 20.37 | 811.68 |
| Random Forest | 0.400 (±) | 18.92 | 775.25 |
| XGBoost | 0.348 | 20.37 | 811.68 |

Evaluation results show that all three models have performance relatively predictive comparable, although ensemble approaches like Random Forest are a bit more superior in matter accuracy (MAE and MSE are higher) low). Fixed linear regression model own high interpretability and provides clarity connection between variables, but his abilities For catch interaction complex limited. Temporary that, XGBoost is known superior in various competition data prediction is not show superiority significant in the context of this data, it is possible due to size relative sample small and limited predictor. With Thus, the selection of the best model No only based on metric accuracy, but also consider interpretability and needs application. Random Forest can chosen as a prediction model main Because offer balance between accuracy, durability against overfitting, and convenience interpretation through feature importance analysis . Findings This give contribution to practice evaluation learning data- based in schools basis, where the use of Machine Learning can be help in develop learning strategies adaptive and based predictions, with consider complexity variables learning and profile student

## 4. Discussion

The results of this study provide an empirical overview of the comparative effectiveness of Game-Based Learning (GBL) and Problem-Based Learning (PBL) at the elementary education level based on secondary data analysis. Consistently, the findings show that PBL produces higher learning outcomes than GBL in various learning contexts. These findings reinforce previous research results which state that learning that emphasizes problem solving, reflection, and argumentation tends to support deeper conceptual understanding.

The superiority of PBL in this study can be explained through its structural and pedagogical characteristics. PBL encourages students to engage in contextual and open-ended problem-solving processes, thus requiring continuous analysis, evaluation, and reflection skills [16, 19, 20]. This process allows students to construct knowledge in a more meaningful way, especially when learning is carried out over an adequate duration of intervention. Thus, PBL provides more space for the development of higher-order thinking skills compared to approaches that focus solely on game activities [21-23].

On the other hand, GBL shows its main strength in terms of student engagement and motivation to learn. Interactive and visual game elements can increase student interest and participation, especially in exploratory and practical materials [24-26]. However, without adequate reflective structure and cognitive scaffolding, this increase in motivation is not always followed by a significant increase in learning outcomes [27-29]. This may explain why the learning outcomes in GBL in this study were relatively lower than those in PBL.

The analysis results also emphasize the importance of contextual factors in determining learning effectiveness. Students' prior knowledge and the duration of the learning intervention emerged as factors that significantly influenced learning outcomes, regardless of the learning model used [30]. These findings indicate that the effectiveness of GBL and PBL cannot be separated from the characteristics of the students and the design of the learning implementation [3, 31, 32]. Interventions with longer durations provide opportunities for students to engage cognitively and reflectively in the learning process.

The data-driven approach and the use of machine learning techniques in this study serve as exploratory tools to identify important patterns and predictors in the data [15, 33]. These techniques are not intended to replace conventional statistical analysis, but to strengthen the interpretation of results and provide additional perspectives on the factors that influence learning outcomes. This approach demonstrates the potential of secondary data analysis in producing cross-context empirical syntheses of learning.

Despite providing broader insights, this study has limitations that need to be considered. The data analyzed depends on the quality of reporting in the source articles, as well as variations in learning designs and assessment instruments used in each study. Therefore, the results of this study should be interpreted as comparative findings based on secondary data and are not intended to draw direct causal conclusions [34, 35].

Overall, the results of this study indicate that both GBL and PBL have a positive contribution to learning in elementary schools, but PBL shows greater effectiveness in supporting improved learning outcomes. These findings have implications for educators and learning designers in selecting and designing learning models that are appropriate for learning objectives and student characteristics. In addition, the results of this study can also serve as a conceptual basis for further research exploring the development of hybrid learning models empirically through direct implementation in the classroom.

## 5. Final Considerations

This study aims to compare the effectiveness of Game-Based Learning (GBL) and Problem-Based Learning (PBL) at the elementary education level through a secondary data-based approach. Based on the results of quantitative analysis of data extracted from empirical articles, it can be concluded that both learning models contribute positively to student learning outcomes. However, PBL shows higher learning outcomes than GBL, both statistically and practically. In addition to the differences between the two learning models, this study also found that contextual factors, particularly students' prior knowledge and the duration of the learning intervention, play an important role in influencing learning outcomes. These findings indicate that the effectiveness of a learning model is not only determined by the pedagogical approach used, but also by the characteristics of the learners and the design of the learning implementation. The secondary data analysis approach used in this study allows for a broader comparison across learning contexts than single-class studies. By combining conventional statistical analysis and machine learning techniques as exploratory tools, this study provides additional insights into the factors that contribute to successful learning in elementary schools. However, the findings of this study need to be interpreted taking into account the limitations of secondary data and variations in the quality of reporting in the source studies. Overall, the results of this study have important implications for learning practices and educational research. For educators and curriculum designers, these findings can serve as a basis for selecting and designing learning models that suit student needs and learning contexts. For researchers, this study demonstrates the potential use of secondary data analysis and data-based approaches in enriching comparative learning studies. Further research is recommended to combine this approach with direct empirical studies in the classroom to gain a more comprehensive understanding of the effectiveness of various learning models.

## References

[1] T. Liu, M. Oubibi, Y. Zhou, and A. Fute, "Research on online teachers' training based on the gamification design: A survey analysis of primary and secondary school teachers," *Heliyon*, vol. 9, no. 4, p. e15053, 2023.

[2] A. Afandi, S. Sajidan, M. Akhyar, and N. Suryani, "Development frameworks of the Indonesian partnership 21st-century skills standards for prospective science teachers: A Delphi Study," *Jurnal Pendidikan IPA Indonesia*, vol. 8, no. 1, pp. 89-100, 2019.

[3] A.-C. E. Ding and C.-H. Yu, "Serious game-based learning and learning by making games: Types of game-based pedagogies and student gaming hours impact students' science learning outcomes," *Computers & Education*, vol. 218, p. 105075, 2024. https://doi.org/10.1016/j.compedu.2024.105075

[4] M. Hartt, H. Hosseini, and M. Mostafapour, "Game on: Exploring the effectiveness of game-based learning," *Planning Practice & Research*, vol. 35, no. 5, pp. 589-604, 2020. https://doi.org/10.1080/02697459.2020.1778859

[5] M. Videnovik, A. M. Bogdanova, and V. Trajkovik, "Game-based learning approach in computer science in primary education: A systematic review," *Entertainment Computing*, vol. 48, p. 100616, 2024. https://doi.org/10.1016/j.entcom.2023.100616

[6] Z. Cai, X. Zhang, C. Liu, and J. Zhan, "Effects of digital game-based learning on student's problem-solving ability: A three-level meta-analysis," *Journal of Computer Assisted Learning,* vol. 41, no. 2, p. e70002, 2025. https://doi.org/10.1111/jcal.70002Digital

[7] M. M. K. Hawamdeh and I. Adamu, *Problem-based learning (PBL): A deep approach to learning in the 21st century. In G. Kurubacak-Meric & S. Sisman-Ugur (Eds.), Improving scientific communication for lifelong learners*. Hershey, PA, USA: IGI Global Scientific Publishing, 2021.

[8] W. K. Niswa, A. Agustiningsih, and K. Mahmudi, "Problem-based learning (PBL): An innovation in natural sciences studies to improve critical thinking skills," *Pancaran Pendidikan,* vol. 11, no. 2, pp. 113–122, 2022.

[9] P. Hegade, O. Patil, and A. Shettar, "Effectiveness of game based learning in problem based learning," presented at the Edition of IEEE Delhi Section Flagship Conference (DELCON), 2023.

[10] Q. Guo, H. Jamil, L. Ismail, S. Luo, and Z. Sun, "Effects of problem-based learning on EFL learning: A systematic review," *PloS One,* vol. 19, no. 12, p. e0307819, 2024.

[11] M. A. Almulla, "The effectiveness of the project-based learning (PBL) approach as a way to engage students in learning," *Sage Open,* vol. 10, no. 3, p. 2158244020938702, 2020. https://doi.org/10.1177/2158244020938702

[12] Y. Liu and A. Pásztor, "Effects of problem-based learning instructional intervention on critical thinking in higher education: A meta-analysis," *Thinking Skills and Creativity,* vol. 45, p. 101069, 2022. https://doi.org/10.1016/j.tsc.2022.101069

[13] Y. Li, X. Wang, X.-r. Zhu, Y.-x. Zhu, and J. Sun, "Effectiveness of problem-based learning on the professional communication competencies of nursing students and nurses: A systematic review," *Nurse Education in Practice,* vol. 37, pp. 45-55, 2019. https://doi.org/10.1016/j.nepr.2019.04.015

[14] W. Setyaningrum, L. D. Pratama, and M. B. Ali, "Game-based learning in problem solving method: The effects on students' achievement," *International Journal on Emerging Mathematics Education,* vol. 2, no. 2, p. 157, 2018.

[15] J. S. Kinnebrew, J. R. Segedy, and G. Biswas, "Integrating model-driven and data-driven techniques for analyzing learning behaviors in open-ended learning environments," *IEEE Transactions on Learning Technologies,* vol. 10, no. 2, pp. 140-153, 2015. https://doi.org/10.1109/TLT.2015.2513387

[16] I. M. C. Wibawa, I. W. Susanta, N. N. Parmithi, and I. W. E. Mahendra, "Improving the scientific attitude of elementary school students through problem-based learning," *MIMBAR PGSD Undiksha,* vol. 11, no. 1, pp. 18-23, 2023. https://doi.org/10.23887/jjpgsd.v11i1.61884

[17] L. Fu, Y. Li, Y. Wu, X. Wang, and Y. Jiang, "Low carbon district heating in China in 2025-a district heating mode with low grade waste heat as heat source," *Energy,* vol. 230, p. 120765, 2021.

[18] J. Gu *et al.*, "NTIRE 2022 challenge on perceptual image quality assessment," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022.

[19] A. Budhiman, O. Satifah, and F. P. Sari, "The influence of problem-based learning on elementary school students' interest in science," *Impulse: Journal of Research and Innovation in Physics Education,* vol. 2, no. 1, pp. 55-63, 2022. https://doi.org/10.14421/impulse.2022.21.06

[20] S. Susilawati and T. Supriyatno, "Problem-Based Learning model in improving critical thinking ability of elementary school students," *Advances in Mobile Learning Educational Research,* vol. 3, no. 1, pp. 638-647, 2023. https://doi.org/10.25082/AMLER.2023.01.013

[21] S. Ndiung and S. Menggo, "Problem-based learning analysis in strengthening college students' interpersonal communication skill," *Journal of Education and Learning (EduLearn),* vol. 18, no. 4, pp. 1350-1361, 2024. https://doi.org/10.11591/edulearn.v18i4.21219

[22] J. Kloeg, "Education as an open question: A hermeneutical approach to problem-based learning," *Journal of Problem Based Learning in Higher Education,* vol. 11, no. 1, pp. 79-97, 2023. https://doi.org/10.54337/ojs.jpblhe.v11i1.7373

[23] C.-L. Tsai *et al.*, "Effectiveness of tutor shadowing on faculty development in problem-based learning," *BMC Medical Education,* vol. 22, no. 1, p. 564, 2022. https://doi.org/10.1186/s12909-022-03615-0

[24] A. Christopoulos, S. Mystakidis, J. Kurczaba, M.-J. Laakso, and C. Stylios, "Is immersion in 3D virtual games associated with mathematical ability improvement in game-based learning?," *International Journal of Science and Mathematics Education,* vol. 22, no. 7, pp. 1479-1499, 2024. https://doi.org/10.1007/s10763-023-10440-4

[25] A. Ilhan, "The impact of game-based, modeling, and collaborative learning methods on the achievements, motivations, and visual mathematical literacy perceptions," *SAGE Open,* vol. 11, no. 1, p. 21582440211003567, 2021. https://doi.org/10.1177/21582440211003567

[26] C.-H. Kuo, M.-J. Chen, R. Nababan, and H.-C. She, "Space adventure game-based learning: How games and scaffolds affect eighth graders' physics learning and game immersion," *IEEE Transactions on Learning Technologies,* vol. 17, pp. 229-240, 2023. https://doi.org/10.1109/TLT.2023.3288879

[27] F. Machon *et al.*, "Design of individual simulation games in manufacturing companies for game-based learning," *Procedia CIRP,* vol. 119, pp. 1017-1022, 2023. https://doi.org/10.1016/j.procir.2023.03.145

[28] D. Y. Saputri, M. Indriayu, and Rukayah, "Learning transformation 4.0 with games-based interactive multimedia: Its effectiveness towards elementary school students' cognitive skills," *Pegem Journal of Education and Instruction,* vol. 14, no. 4, 2024. https://doi.org/10.47750/pegegog.14.04.36

[29] A. A. Al-Barakat, R. M. AlAli, O. M. Al-Hassan, and Y. Z. Aboud, "Evaluating the effectiveness of electronic games-based learning in enhancing children's multiplication skills and cognitive achievement," *International Journal of Information and Education Technology,* vol. 15, no. 5, pp. 947-954, 2025. https://doi.org/10.18178/ijiet.2025.15.5.2301

[30] M. M. Roshid and M. Z. Haider, "Teaching 21st-century skills in rural secondary schools: From theory to practice," *Heliyon,* vol. 10, no. 9, p. e30769, 2024. https://doi.org/10.1016/j.heliyon.2024.e30769

[31] S. N. Ajid, D. A. Kusumaningtyas, K. Ratih, and S. Lava, "Strategies for integrating problem-based learning, teaching modules, and formative assessments to enhance learning outcomes and critical thinking skills," *Indonesian Journal on Learning and Advanced Education,* pp. 218-232, 2025.

[32] P. Pimdee, A. Sukkamart, C. Nantha, T. Kantathanawat, and P. Leekitchwatana, "Enhancing Thai student-teacher problem-solving skills and academic achievement through a blended problem-based learning approach in online flipped classrooms," *Heliyon,* vol. 10, no. 7, p. e29172, 2024. https://doi.org/10.1016/j.heliyon.2024.e29172

[33] D. Čubela, A. Rossner, and P. Neis, "Using problem-based learning and gamification as a catalyst for student engagement in data-driven engineering education: A report," *Education Sciences*, vol. 13, no. 12, p. 1223, 2023. https://doi.org/10.3390/educsci13121223

[34] N. Maher, "Perspectives on pedagogical content knowledge in the senior secondary mathematics classroom," 2019. https://consensus.app/papers/perspectives-on-pedagogical-content-knowledge-in-the-maher/412d389876f05dfdb8a194a339c56680/

[35] T. M. Yuliandari, A. Putri, and Y. Rosmansyah, "Digital transformation in secondary schools: A systematic literature review," *IEEE Access*, vol. 11, pp. 90459-90476, 2023. https://doi.org/10.1109/ACCESS.2023.3306603