BIMA
BULLETIN OF INTELLIGENT MACHINES AND ALGORITHMS

Vol. 1 No. 4 May 2026
E-ISSN: 3132-5115
DOI: 10.65780/bima.v1i4.23

Preprocessing and Feature Engineering of Gameplay Logs for Adaptive Mathematics Learning Dataset Construction

Rio Andriyat Krisdiawan[1*], Dede Husen[2], Heri Herwanto[3]

Informatics Engineering, Faculty of Computer Science[1,2,3]
Kuningan University, Kuningan, Indonesia[1,2,3]
http://www.fkom.uniku.ac.id/ [1,2,3]
rioandriyat@uniku.ac.id[1*], dede.husen@uniku.ac.id[2], heri.herwanto@uniku.ac.id[3]

Abstract. This study reports the preprocessing and feature engineering of gameplay logs collected from a first-grade elementary mathematics educational game prototype. The study aimed to transform raw gameplay activity records into structured analytical datasets that can support early performance description and serve as an initial data preparation stage for future adaptive mathematics learning research. A limited trial was conducted with nine first-grade students who played six sequential game levels covering early numeracy topics, including counting, number ordering, number reading, place value, and mixed review. The gameplay logs captured event-based student interactions, including session identity, level, question, selected answer, correctness status, attempt count, help usage, response time, score, and event type. The data processing workflow included data validation, cleaning, anonymization, data type handling, event filtering, feature engineering, dataset aggregation, and descriptive analysis. The preprocessing stage produced 262 clean gameplay log records consisting of 171 answer events, 28 help events, 54 level completion events, and nine game completion events. Feature engineering generated analytical indicators such as level accuracy, average response time, total help usage, average attempt count, performance category, student state, and initial adaptation action. The final outputs were organized into answer-level, level-level, student-level, and adaptive feature datasets. The anonymized dataset and data dictionary are provided as supplementary materials to support reproducibility and future reuse. The results indicate that raw gameplay logs can be converted into structured datasets for early learning analytics and preliminary adaptive learning data preparation, without making claims about learning effectiveness or final adaptive model performance.

Key words: Adaptive Learning Dataset; Data Preprocessing; Educational Game; Feature Engineering; Gameplay Log; Mathematics Learning.

1. INTRODUCTION

Early numeracy is a fundamental component of mathematics learning in elementary education. In the first grade, students begin to develop basic mathematical understanding through activities such as counting objects, recognizing number sequences, reading number symbols, and identifying simple place value. These early concepts are not only important for mathematics achievement but also serve as a foundation for logical reasoning, problem solving, and further mathematical learning. Therefore, mathematics learning at the early elementary level requires instructional media that are concrete, visual, interactive, and appropriate for the cognitive characteristics of young learners.

Digital game-based learning has been widely explored as an approach to support mathematics learning. Educational games can provide structured activities through challenges, feedback, visual representation, repeated practice, and gradual progression. Previous studies have shown that digital games in primary mathematics learning may support student engagement, enrich learning experiences, and provide interactive ways to strengthen mathematical concepts [1], [2], [3], [4]. In the Indonesian context, educational games have also been developed for elementary mathematics learning, including counting activities, Android-based learning media, and mathematics games designed to support critical thinking [5], [6], [7].

Beyond their role as instructional media, educational games can also function as sources of learning activity data. Each student interaction during gameplay, such as answering questions, requesting help, making attempts, completing levels, and finishing the game, can be recorded as a gameplay log. These logs provide process-oriented information that is richer than final scores alone. Gameplay logs can capture how students interact with learning tasks,

Vol. 1 No. 4 May 2026
E-ISSN: 3132-5115
DOI: 10.65780/bima.v1i4.23

how quickly they respond, whether they need support, and how they progress across game levels. Such data are relevant to learning analytics because they allow student learning behavior to be examined through digital activity traces [8], [9].

In game-based learning environments, gameplay logs can be transformed into analytical indicators such as answer accuracy, response time, number of attempts, help usage, score, and level completion status. These indicators can support early student performance profiling and may serve as input for adaptive learning development. In game-based assessment and stealth assessment, student actions embedded in gameplay can be treated as evidence of learning-related behavior, allowing learning information to be collected without relying solely on external tests [9], [10]. This perspective is particularly relevant for adaptive learning because an adaptive system requires structured information about student conditions before providing differentiated learning support.

However, raw gameplay logs cannot be used directly as research datasets. Game activity data are commonly event-based and heterogeneous. A single log file may contain different event types, such as answer events, help events, level completion events, and game completion events. Each event type has a different analytical meaning and may contain different patterns of missing values. For example, an empty answer field in a help or level completion event should not be interpreted in the same way as an empty answer field in an answer event. In addition, some student responses may appear as numbers, while others may be represented as text phrases, such as number sequences or place-value descriptions. These characteristics require systematic preprocessing before the data can be used for analysis.

Feature engineering is also needed to convert raw gameplay events into meaningful learning indicators. Features such as level accuracy, average response time, total help usage, average attempt count, performance category, student state, and initial adaptation action are not always available directly from the raw log structure. These features must be constructed through validation, cleaning, transformation, contextual interpretation, and aggregation. Without these stages, gameplay logs remain technical records of system activity rather than analytical datasets that can support learning analytics or adaptive model development.

Previous studies have discussed educational games, learning analytics, game-based assessment, stealth assessment, and adaptive support in game-based learning. Nevertheless, many of these studies focus on learning effectiveness, media development, student experience, or adaptive model implementation. Fewer studies explain in detail how gameplay logs from early-grade mathematics educational games can be processed into structured analytical datasets through preprocessing and feature engineering. This gap is important because adaptive learning development requires reliable data preparation before any adaptive or predictive model can be meaningfully designed and evaluated.

This study addresses that gap by presenting the preprocessing and feature engineering of gameplay logs collected from a first-grade elementary mathematics educational game prototype. The study focuses on transforming raw event-based gameplay records into structured analytical datasets. The contribution of this study lies in the construction of answer-level, level-level, student-level, and adaptive feature datasets that can support early learning data analysis and serve as a preliminary data preparation stage for future adaptive mathematics learning research.

This article is deliberately limited to gameplay log dataset construction. It does not evaluate learning effectiveness, does not apply a pretest-posttest design, and does not assess the performance of a final adaptive or Q-Learning model. Specifically, this study addresses four research questions: first, how the gameplay log structure is generated by the first-grade mathematics educational game prototype; second, how preprocessing is applied to produce clean and structured gameplay logs; third, how feature engineering is used to construct analytical features from gameplay data; and fourth, what forms of analytical datasets are produced as an initial foundation for future adaptive learning model development.

2. RELATED WORK

Previous studies have examined educational games, learning analytics, game-based assessment, and adaptive support from different perspectives. However, the specific issue addressed in this study is not the effectiveness of educational games or the performance of an adaptive model, but the construction of structured analytical datasets from gameplay logs. Table 1 summarizes relevant previous studies and clarifies the research gap addressed in this article.

TABLE 1. Previous Studies and Research Gap

Previous Study	Research Focus	Method / Context	Limitation Related to This Study	Gap Addressed in This Article
Dan et al. [1]	Digital game-based learning in primary mathematics education	Systematic literature review on digital game-based learning in elementary mathematics	Focuses on research trends, learning topics, and learning outcomes, not on gameplay log dataset construction	This article focuses on constructing a structured gameplay log dataset for first-grade mathematics learning
Debrenti [2]	Game-based learning experience	Study of game-based mathemati	Emphasizes learning experience and	This article prepares structured gameplay

BIMA
BULLETIN OF INTELLIGENT MACHINES AND ALGORITHMS
Vol. 1 No. 4 May 2026
E-ISSN: 3132-5115
DOI: 10.65780/bima.v1i4.23

Previous Study	Research Focus	Method / Context	Limitation Related to This Study	Gap Addressed in This Article
	s in primary mathematics	cs learning experiences in primary education	classroom implementation rather than event-based data preprocessing	data generated from a mathematics educational game
Gui et al. [3]	Effectiveness of digital educational games in STEM learning	Meta-analysis of digital educational games and game design	Focuses on learning effectiveness and game design effects, not data engineering from gameplay logs	This article shifts the focus from effectiveness evaluation to preprocessing and feature engineering of game activity data
Hidayat et al. [4]	Online game-based learning in mathematics education	Systematic review of online game-based mathematics learning	Discusses online game-based learning trends, not early-grade gameplay log structure	This article positions gameplay logs as data sources for early numeracy learning analytics
Banihashem et al. [8]	Learning analytics for online game-based learning	Review and conceptual discussion of learning analytics in game-based environments	Provides a broad learning analytics perspective, but not a specific preprocessing workflow for grade-1 mathematics logs	This article applies the learning analytics perspective to construct answer-level, level-level, and student-level datasets
Lu et al. [9]	Game-based learning prediction model construction	Stealth assessment and prediction model pipeline in game-based learning	Focuses on prediction and stealth assessment validation, not initial dataset construction in early numeracy games	This article prepares analytical datasets that may support future prediction or adaptive learning models
Shute et al. [10]	Stealth assessment, adaptivity, and learning supports in	Educational game-based assessment and adaptive learning support	Focuses on assessment validity and adaptivity, not preprocessing of	This article provides the data preparation stage before future adaptive
	educational games		classroom gameplay logs	learning implementation
Chen et al. [11]	Adaptive scaffolding and engagement in digital game-based learning	Adaptive scaffolding in interactive digital game environments	Focuses on scaffolding and engagement effects, not construction of adaptive features from raw logs	This article constructs initial features such as help usage, student state, and initial adaptation action
Rivaldi and Kurniawan [5]	Educational counting game for first-grade students	Development of a grade-1 counting educational game	Focuses on game development, not gameplay log preprocessing and feature engineering	This article extends the grade-1 mathematics game context toward structured gameplay log dataset construction
Damarjati and Miatun [7]	Android-based educational game for mathematics learning	Educational game development for mathematics and critical thinking	Does not discuss event-based gameplay logs or adaptive dataset features	This article contributes a data engineering perspective for educational game research
Krisdiawan et al. [12]	UI/UX design for educational AR card game	User-centered design in educational AR card game	Focuses on interface and user experience, not gameplay log analytics	This article adds a gameplay data construction layer to educational game research

Digital game-based learning has been widely examined in mathematics education because it combines instructional content with interaction, challenge, feedback, and visual representation. In primary mathematics learning, educational games can support students through gradual tasks, repeated practice, and contextual learning activities [1], [2]. Broader reviews and meta-analyses also indicate that the educational value of digital games is influenced by the alignment between game design, learning objectives, instructional content, and the nature of student interaction [3], [13], [14].

National studies have also shown that educational games can be developed for elementary mathematics contexts, including counting activities, Android-based learning media, and mathematics learning media designed to

Vol. 1 No. 4 May 2026
E-ISSN: 3132-5115
DOI: 10.65780/bima.v1i4.23

support critical thinking [5], [6], [7]. These studies provide an important pedagogical foundation, but most of them focus on game development, user experience, or learning outcomes rather than the preparation of gameplay activity data.

Gameplay logs offer a different perspective because they capture process-oriented data generated during interaction with a game-based learning environment. A gameplay log may record answer events, help requests, response times, attempts, scores, level completion, and game completion. Such data can provide more detailed information than final scores because they describe how students interact with learning tasks during gameplay. In learning analytics, digital activity traces are used to understand learning behavior, identify performance patterns, and support data-informed educational decisions [8], [15], [16]. Therefore, gameplay logs can be treated as a potential source of learning analytics data, provided that they are processed into a structured and interpretable format.

Learning analytics in game-based learning requires more than collecting raw activity records. Raw logs need to be transformed into features that can represent student behavior and performance. Banihashem et al. [8] emphasized the role of learning analytics in understanding online game-based learning through student activity data. Daoudi [15] also highlighted that learning analytics in serious games can support the interpretation of user behavior and educational interaction in formal learning contexts. In addition, Lu et al. [9] described the importance of connecting gameplay data, feature processing, and model interpretation in game-based learning prediction pipelines. These perspectives suggest that preprocessing and feature engineering are necessary stages before gameplay logs can be used for prediction, assessment, or adaptive learning purposes.

Game-based assessment and stealth assessment further support the use of gameplay activity as learning evidence. In this perspective, student actions embedded in gameplay can be interpreted as indicators of performance, engagement, or learning-related behavior [10]. This approach is relevant to adaptive learning because adaptive systems require structured information about student conditions before delivering differentiated support. However, this study does not validate an assessment model or evaluate adaptive learning performance. Instead, it focuses on the earlier stage: preparing clean and structured datasets from event-based gameplay logs.

Adaptive game-based learning and adaptive scaffolding studies indicate that student support should be informed by interaction data. Information such as response accuracy, response time, number of attempts, help usage, and level completion can help describe student conditions more comprehensively than final scores alone. Chen et al. [11] discussed the relationship between adaptive scaffolding and student engagement in digital game-based learning, while Faber et al. [17] examined adaptive scaffolding in relation to performance, cognitive load, and engagement. Chiotaki et al. [18] also showed that adaptive game-based learning is closely related to the personalization of content and learning experience. These studies reinforce the need for structured gameplay data, but they do not specifically address the preprocessing and feature engineering of early-grade mathematics gameplay logs.

From a data engineering perspective, gameplay logs are challenging because they are usually event-based and heterogeneous. One log file may include multiple event types, and each event type may contain different values and meanings. For example, an empty selected answer field in a help or level completion event should not be interpreted as the same type of missing data as an empty selected answer field in an answer event. Therefore, preprocessing must consider the semantic meaning of each event type. This includes validating the data structure, checking duplicate records, interpreting missing values according to event context, anonymizing student identity, handling data types, and filtering events for specific analytical purposes.

Feature engineering is the next essential stage in transforming raw gameplay records into analytical datasets. Features such as level accuracy, average response time, total help usage, average attempt count, performance category, student state, and initial adaptation action are not directly available in raw event logs. They must be derived through calculation, transformation, and aggregation. Lu et al. [19] showed that feature generation from game telemetry and game metrics is an important part of serious game analytics, while Calvo-Morata et al. [20] demonstrated how learning analytics can guide serious game development through the interpretation of gameplay data. In this study, feature engineering is used to construct answer-level, level-level, student-level, and adaptive feature datasets from a first-grade mathematics educational game prototype.

Based on the reviewed literature, this article is positioned at the data preparation stage of adaptive mathematics learning research. Previous studies have provided important insights into educational games, learning analytics, stealth assessment, adaptive scaffolding, and serious game analytics. However, there remains a specific need to describe how gameplay logs from early-grade mathematics educational games can be transformed into structured analytical datasets. This study addresses that need by presenting a preprocessing and feature engineering workflow for gameplay log dataset construction. The scope is deliberately limited to dataset construction and does not include learning effectiveness evaluation, reward analysis, Q-value updates, or final Q-Learning model performance.

3. METHODS

This study applied a gameplay log-based dataset construction approach. The main objective was to transform raw gameplay activity records into clean,

BIMA
BULLETIN OF INTELLIGENT MACHINES AND ALGORITHMS
Vol. 1 No. 4 May 2026
E-ISSN: 3132-5115
DOI: 10.65780/bima.v1i4.23

structured, and analyzable datasets that can support early learning data analysis and serve as an initial data preparation stage for future adaptive mathematics learning research. This study was not designed as an experimental study to evaluate learning effectiveness, did not apply a pretest-posttest design, and did not assess the performance of a final adaptive or Q-Learning model. Instead, the methodological focus was placed on data validation, data cleaning, anonymization, event filtering, feature engineering, dataset aggregation, and descriptive analysis. This approach is consistent with the role of learning analytics in game-based learning, where activity data must first be prepared and structured before being used for interpretation, prediction, or adaptive learning development [8], [9].

3.1. Educational Game Prototype and Data Source.

The data source was a first-grade elementary mathematics educational game prototype designed for early numeracy learning. The game consisted of six sequential levels, each representing a different learning mission. The learning content included counting objects, identifying number order, reading numbers, understanding simple place value, and reviewing mixed numeracy concepts. In this study, the game was positioned as a data-generating environment that produced gameplay logs, not as an intervention medium for evaluating learning effectiveness. The structure of the game levels is shown in Table 2.

TABLE 2. Structure of the Mathematics Educational Game Prototype

Level	Mission Name	Material Type	Learning Focus
1	Counting Fish Mission	Counting	Counting simple objects
2	Number Garden Mission	Counting	Counting objects in varied contexts
3	Number Stone Mission	Ordering	Identifying number order or missing numbers
4	Number Forest Mission	Number Reading	Recognizing and reading numbers
5	Tens Basket Mission	Place Value	Understanding tens and ones
6	Number House 20 Mission	Mixed Review	Reviewing early numeracy concepts

A limited trial was conducted with nine first-grade elementary students. Each student played the game from Level 1 to Level 6. The dataset analyzed in this study consisted of gameplay logs recorded during student interaction with the game. Because the trial involved a small number of students, the data were treated as an initial dataset for validating the log structure, preprocessing workflow, and feature engineering procedure, rather than as a basis for broad generalization.

Student identities were anonymized before analysis. No personally identifiable information was presented in the dataset or in this manuscript. The anonymization process was conducted to ensure that the gameplay logs could be analyzed while preserving student privacy.

3.2. Gameplay Log Structure

The gameplay log was recorded automatically while students interacted with the game. Each row in the log represented one gameplay event. The recorded event types consisted of answer, help, level_complete, and game_complete. The answer event represented a student response to a question. The help event represented the use of assistance during gameplay. The level_complete event indicated that a student had completed one game level, while the game_complete event indicated that the student had completed all six levels.

The gameplay log was designed to capture key information related to the learning process, including session identity, game level, question, selected answer, correctness status, attempt count, help usage, response time, score, accumulated score, and event type. This structure enabled the data to be analyzed at multiple levels, including answer-level, level-level, and student-level analysis. The data dictionary is presented in Table 3.

TABLE 3. Gameplay Log Data Dictionary

Column	Description	Data Type
timestamp	Time when the event was recorded	Datetime
session_id	Unique gameplay session identifier	String
student_id	Anonymized student identifier	String
level	Game level number	Integer
level_name	Name of the game mission	String
material_type	Mathematics material type	String
question_id	Unique question identifier	String
question_text	Question displayed in the game	String
selected_answer	Student's selected answer	String/Numeric
correct_answer	Correct answer	String/Numeric
is_correct	Correctness status	Integer
attempt_count	Number of answer attempts	Integer
help_used	Help usage status	Integer
response_time_seconds	Response time in seconds	Numeric
score	Score obtained from the event	Integer
total_score	Accumulated score	Integer
event_type	Gameplay event type	String

The selected_answer and correct_answer fields were retained as text or mixed text-numeric fields because some student responses were not purely numeric. Several answers could appear as phrases, such as place-value descriptions or number sequences. Therefore, these fields were not forced into numeric format in order to preserve the original meaning of student responses.

Vol. 1 No. 4 May 2026
E-ISSN: 3132-5115
DOI: 10.65780/bima.v1i4.23

3.3. Data Preprocessing Procedure

Data preprocessing was conducted to ensure that the raw gameplay logs could be used as a research dataset. The preprocessing procedure consisted of data validation, data cleaning, anonymization, data type handling, missing value interpretation, and event filtering. The overall workflow is shown in Figure. 1.

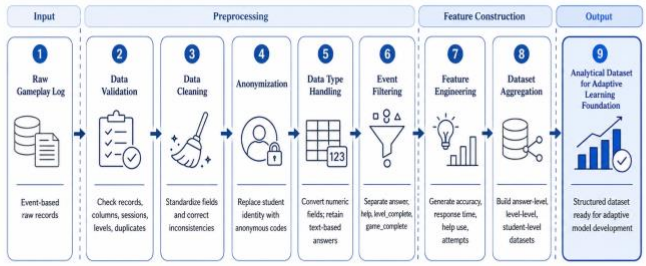

Fig.1. Gameplay Log Preprocessing and Feature Engineering Workflow

The first step was data validation. This step examined the number of records, columns, unique students, gameplay sessions, recorded levels, event types, duplicate records, and game completion status. The purpose of this validation was to determine whether the raw logs were consistent with the designed logging structure of the game prototype.

The second step was data cleaning. This process included standardizing column names, checking duplicate records, and examining value consistency across columns. Student identities were then anonymized by replacing original identifiers with anonymous student codes. This procedure ensured that the dataset could be analyzed without exposing student identity.

The third step was data type handling. Numeric fields such as level, is_correct, attempt_count, help_used, response_time_seconds, score, and total_score were converted into appropriate numeric formats. Meanwhile, selected_answer and correct_answer were retained as text or mixed-format fields because several answers contained phrases rather than only numbers.

The fourth step was contextual missing value interpretation. Empty values were interpreted based on event type. For example, an empty answer field in help, level_complete, or game_complete events was not treated as a data error because those events did not represent student answer submissions. This contextual interpretation was necessary because gameplay logs are event-based datasets in which each event type has a different analytical meaning.

The fifth step was event filtering. The gameplay log was separated into event-based subsets, including answer logs, help logs, level completion logs, and game completion logs. This separation was necessary because each event type supported a different analytical purpose. Answer logs were used for student response analysis, help logs for help usage analysis, level completion logs for level progression validation, and game completion logs for validating full game completion. The preprocessing procedure is summarized in Table 4.

TABLE 4. Preprocessing Procedure

Stage	Process	Output
Data validation	Checking records, columns, sessions, levels, event types, duplicates, and completion status	Raw data validity status
Data cleaning	Standardizing column names and checking inconsistent values	Clean gameplay log
Anonymization	Replacing student identity with anonymous codes	Ethical research dataset
Data type handling	Converting numeric columns and retaining text-based answers	Consistent data format
Missing value interpretation	Interpreting empty values based on event type	Context-aware data quality
Event filtering	Separating answer, help, level completion, and game completion events	Event-based subsets

3.4. Feature Engineering and Dataset Aggregation.

Feature engineering was conducted to transform gameplay logs into analytical indicators that could describe early student performance. This step was necessary because raw logs only recorded gameplay events, while learning data analysis requires derived indicators with clearer analytical meaning. In game-based learning analytics, feature engineering acts as a bridge between activity logs or game telemetry and datasets that can be used for performance analysis, learning analytics, and future model development [9], [19].

The engineered features included accuracy_level, avg_response_time, total_help_used, avg_attempt_count, total_accuracy, performance_category, student_state, and adaptation_action. Accuracy-related features were used to describe student correctness. Response time was used to represent the speed of task completion. Help usage and attempt count were used to identify possible scaffolding needs. The student_state and adaptation_action variables were constructed as preliminary rule-based features to prepare an adaptive feature dataset for future development.

The student_state and adaptation_action variables were not treated as the output of a final artificial intelligence model. They were preliminary rule-based labels derived from level accuracy and help usage. Therefore, these variables should be interpreted as initial adaptive features for dataset preparation, not as evidence of adaptive model performance. The engineered features are summarized in Table 5.

Vol. 1 No. 4 May 2026
E-ISSN: 3132-5115
DOI: 10.65780/bima.v1i4.23

TABLE 5. Engineered Features

Feature	Construction Technique	Analytical Function
accuracy_level	Mean of correctness status in each level	Measuring level mastery
avg_response_time	Mean response time per level/student	Measuring response speed
total_help_used	Sum of help usage per level/student	Identifying scaffolding need
avg_attempt_count	Mean number of attempts	Identifying answering effort
total_accuracy	Mean correctness across all answers	Measuring overall performance
performance_category	Categorization based on total accuracy	Summarizing student performance
student_state	Rule-based state from accuracy and help usage	Initial adaptive state feature
adaptation_action	Rule-based action from student state	Initial adaptation feature

The initial rule-based mapping for student_state and adaptation_action was designed based on a simple pedagogical interpretation of mastery, stability, and learning support needs. In this study, level accuracy was used as an indicator of task mastery, while help usage was used as an indicator of scaffolding dependency. A student with high accuracy and no help usage was interpreted as showing independent mastery; therefore, the initial adaptive action was to increase the level of challenge. A student with moderate accuracy was interpreted as having adequate but still developing performance; therefore, the initial action was to maintain the current level. A student with low accuracy was interpreted as needing additional reinforcement; therefore, the initial action was to decrease the challenge or provide reinforcement. The thresholds were used as preliminary rule-based labels for dataset construction and were not intended to represent a validated adaptive model. The initial rule-based mapping used to construct student_state and adaptation_action is shown in Table 6.

TABLE 6. Initial Rule-Based Mapping for Student State and Adaptation Action

Condition	Student State	Initial Adaptation Action	Pedagogical Rationale
accuracy_level ≥ 0.80 and total_help_used = 0	Mastered	Increase	High accuracy without help indicates independent mastery; higher challenge can be introduced
accuracy_level ≥ 0.60	Adequate	Maintain	Moderate performance indicates sufficient progress but still requires stabilization
accuracy_level < 0.60	Struggling	Decrease/Reinforcement	Low accuracy indicates the need for reinforcement or reduced difficulty

After feature construction, the gameplay logs were aggregated into several units of analysis. The answer-level dataset represented student responses to individual questions. The level-level dataset represented student performance in each level. The student-level dataset summarized each student's overall gameplay performance. In addition, the adaptive feature dataset was prepared to store preliminary adaptive features for future adaptive learning model development. The resulting analytical datasets are summarized in Table 7.

TABLE 7. Analytical Dataset Outputs

Dataset	Unit of Analysis	Function
Answer-level dataset	One row represents one student response to one question	Analyzing answer accuracy, response time, attempts, and help use
Level-level dataset	One row represents one student's performance in one level	Analyzing level performance and student state
Student-level dataset	One row represents one student summary	Analyzing overall student performance
Adaptive feature dataset	One row represents initial adaptive features	Providing preliminary input for future adaptive model development

3.5. Descriptive Data Analysis.

The data analysis in this study was limited to descriptive statistics. The purpose of the analysis was not to evaluate learning effectiveness, but to examine data completeness, event distribution, gameplay log consistency, and the structure of the resulting analytical datasets. The analyzed indicators included the number of records, number of students, number of sessions, event distribution, number of logs per level, level accuracy, average response time, help usage, student state distribution, and adaptation action distribution.

Descriptive analysis was also used to examine whether the gameplay logs could be transformed into multiple analytical datasets. The results of this analysis provided evidence of whether the data structure produced by the prototype was sufficient for initial dataset construction and future adaptive learning model development. This study did not analyze learning improvement, did not calculate N-Gain, did not conduct pretest-posttest statistical testing, and did not evaluate reward values, Q-values, or Q-Learning performance. These boundaries were applied to keep the

Vol. 1 No. 4 May 2026
E-ISSN: 3132-5115
DOI: 10.65780/bima.v1i4.23

article focused on preprocessing, feature engineering, and initial gameplay log dataset construction.

4. RESULTS AND DISCUSSIONS

4.1. Raw Gameplay Log Validation and Event Distribution.

The initial validation showed that the gameplay logs generated by the first-grade mathematics educational game prototype had a consistent structure for further preprocessing. The raw dataset consisted of 262 gameplay log records with 20 initial columns. The data were obtained from nine first-grade students, each represented by one unique gameplay session. All sessions recorded gameplay activities from Level 1 to Level 6, and all students had a game_complete event. No duplicate records were found in the raw dataset. These results indicate that the logging mechanism was able to capture the essential gameplay activities required for initial dataset construction in the limited trial.

TABLE 8. Raw Gameplay Log Validation Summary

Validation Aspect	Result
Raw log records	262
Initial columns	20
Unique students	9
Unique sessions	9
Recorded levels	6
Level range	1–6
Duplicate records	0
Students completing the game	9
Game completion events	9

Table 8 shows that the raw gameplay logs had sufficient structural completeness to be processed into analytical datasets. The equal number of students and sessions indicates that each student had a separate gameplay record. The presence of six recorded levels in the sessions also allowed the data to be analyzed at the level of game missions. In addition, the absence of duplicate records strengthened the initial data quality because each recorded event could be treated as a unique gameplay activity.

The distribution of gameplay events is shown in Table 9. The answer event was the most frequent event, with 171 records. This event became the main basis for constructing the answer-level dataset. The help event appeared in 28 records and was used to identify help usage during gameplay. The level_complete event appeared in 54 records, corresponding to nine students completing six levels. The game_complete event appeared in nine records, indicating that all students completed the game.

TABLE 9. Distribution of Gameplay Log Events

Event Type	Number of Records	Analytical Role
answer	171	Student response analysis
help	28	Help usage analysis
level_complete	54	Level completion validation
game_complete	9	Game completion validation
Total	262	Full gameplay activity records

The event distribution indicates that the gameplay log did not only capture student answers but also recorded learning-related interactions during gameplay. This is important because learning data from educational games should not be limited to final scores. Help usage, response time, attempts, and level completion provide additional context for understanding how students interact with learning tasks. Therefore, the event structure of this dataset supports the use of gameplay logs as process-oriented data in game-based learning analytics [7], [8].

4.2. Data Cleaning and Missing Value Interpretation

Data cleaning was conducted to ensure that each column and value in the gameplay log could be interpreted appropriately. No duplicate records were identified, so no record removal was required for duplication. The cleaning process focused on standardizing column names, checking empty values, anonymizing student identity, and ensuring consistency in data types.

One important aspect of preprocessing gameplay logs is the interpretation of missing values based on event type. Empty values in the selected_answer field cannot be automatically treated as problematic missing data because not all events represent answer submissions. In the answer event, no empty value was found in the selected_answer field. In contrast, empty values in help, level_complete, and game_complete events were acceptable because these events do not record student answers.

TABLE 10. Missing Value Interpretation by Event Type

Event Type	Number of Records	Empty Selected Answer	Interpretation
answer	171	0	No missing answer data
help	28	28	Acceptable because it is not an answer event
level_complete	54	54	Acceptable because it records level completion
game_complete	9	9	Acceptable because it records game completion

Table 10 shows that missing value interpretation in gameplay logs must be context-aware. If all empty fields were treated as data errors, non-answer events could be incorrectly interpreted as incomplete records. Therefore, the preprocessing procedure in this study considered the semantic meaning of each event type within the game flow. This approach is important for maintaining the validity of event-based datasets, especially when the dataset contains multiple types of gameplay activities.

BIMA
BULLETIN OF INTELLIGENT MACHINES AND ALGORITHMS
Vol. 1 No. 4 May 2026
E-ISSN: 3132-5115
DOI: 10.65780/bima.v1i4.23

In addition, the selected_answer and correct_answer fields were retained as text or mixed text-numeric data. This decision was made because some student responses were not purely numeric. Several answers could appear as phrases, such as place-value descriptions or number sequences. By retaining the original response format, the meaning of student answers was preserved and not lost through inappropriate numeric conversion.

4.3. Analytical Dataset Construction.

After validation, cleaning, anonymization, and event filtering, the gameplay logs were transformed into several analytical datasets. Dataset construction was conducted to make the data usable at different levels of analysis, including full event records, student answers, help usage, level completion, game completion, answer-level performance, level-level performance, and student-level summaries.

TABLE 11. Analytical Dataset Outputs

Dataset Output	Analytical Function
Clean gameplay log	Validated and anonymized full gameplay activity records
Answer log	Subset for student answer analysis
Help log	Subset for help usage analysis
Level completion log	Subset for validating level completion
Game completion log	Subset for validating full game completion
Answer-level dataset	One row represents one student response
Level-level dataset	One row represents one student performance in one level
Student-level dataset	One row represents one student summary
Adaptive feature dataset	Initial dataset containing student state and adaptation action

Table 11 shows that one gameplay log source can be processed into several datasets with different analytical functions. The answer-level dataset enables analysis of student responses to each question. The level-level dataset enables analysis of student performance in each mission or level. The student-level dataset provides an overall summary of each student's gameplay performance. Meanwhile, the adaptive feature dataset was prepared as an initial dataset containing student_state and adaptation_action.

The construction of multiple analytical datasets is the main contribution of this article. Raw gameplay logs that initially consisted of event records were transformed into more structured analytical datasets. With this structure, the data can be used for initial descriptive analysis and may support future adaptive learning model development. However, the adaptive feature dataset in this article should be interpreted only as a preliminary rule-based feature set, not as the output of a final artificial intelligence model.

4.4. Student-Level and Level-Level Descriptive Results

Student-level descriptive analysis was conducted to obtain an initial overview of student performance based on the constructed dataset. The aggregation results showed that there were 171 answer records, with 161 correct answers. Therefore, the overall answer accuracy was approximately 94.2%. All nine students completed the six game levels, and the dominant performance category was high based on total accuracy.

TABLE 12. Student-Level Dataset Summary

Indicator	Result
Total answer records	171
Total correct answers	161
Overall answer accuracy	94.2%
Unique students	9
Students completing all levels	9
Dominant performance category	High

The results in Table 12 indicate that the prototype captured the selected performance-related indicators during the limited trial. However, the high accuracy should be interpreted carefully. In this article, accuracy is not used to conclude learning effectiveness. Instead, it indicates that the dataset successfully recorded student responses and that the initial level of task difficulty in the prototype may have been manageable for most students. This finding provides input for future development, particularly the need to enrich the variation of difficulty levels.

Level-level analysis was conducted to examine how the level-level dataset represented student performance across game missions. The indicators included average accuracy, average response time, and total help usage. The summary of level-level performance is presented in Table 13.

TABLE 13. Level-Level Performance Summary

Level	Mission Name	Average Accuracy	Average Response Time	Total Help
1	Counting Fish Mission	94.4%	7.91	1
2	Number Garden Mission	94.4%	8.56	4
3	Number Stone Mission	97.2%	7.15	9
4	Number Forest Mission	94.4%	5.02	2
5	Tens Basket Mission	94.4%	7.06	7
6	Number House 20 Mission	97.2%	6.03	6

Table 13 shows that the average accuracy in all levels was high. Levels 3 and 6 had the highest average accuracy, both at 97.2%. However, help usage did not always follow the same pattern as accuracy. Level 3 recorded the highest help usage, with nine help events, even though its accuracy was also high. This result suggests that correctness alone is not sufficient to describe the student learning process during

Vol. 1 No. 4 May 2026
E-ISSN: 3132-5115
DOI: 10.65780/bima.v1i4.23

gameplay. Help usage and response time need to be considered together with accuracy to provide a more complete description of student interaction with the game. These findings support the need for feature engineering in gameplay log processing. If the analysis relies only on score or accuracy, variations in student learning behavior may remain hidden. By constructing features such as avg_response_time, total_help_used, and avg_attempt_count, the dataset can provide richer information about how students complete learning tasks. This is consistent with learning analytics in game-based learning, which emphasizes the importance of activity data for understanding learning processes more closely [8], [19].

4.5. Initial Adaptive Feature Distribution.

Feature engineering also produced initial features intended for future adaptive model development. These features were student_state and adaptation_action. Both were constructed using an initial rule-based mapping that combined level accuracy and help usage. In this study, student_state was categorized into Mastered, Adequate, and Struggling. Meanwhile, adaptation_action was categorized into Increase, Maintain, and Decrease/Reinforcement.

TABLE 14. Initial Adaptive Feature Distribution

Feature Category	Number of Records
Mastered	29
Adequate	24
Struggling	1
Increase	29
Maintain	24
Decrease/Reinforcement	1

The distribution in Table 14 shows that most level-level records were categorized as Mastered or Adequate. Only one record was categorized as Struggling. This pattern is consistent with the relatively high level accuracy observed in the dataset. However, the presence of Adequate and Struggling categories indicates that the initial features did not produce a single dominant label only; they were still able to capture variation in student performance based on the combination of accuracy and help usage.

The student_state and adaptation_action variables should be interpreted strictly as initial rule-based features. They do not represent adaptive model performance and should not be treated as the final output of an artificial intelligence system. Their primary function is to prepare a dataset structure that can support future adaptive model development. Therefore, the results should not be interpreted as evidence of Q-Learning performance or final adaptive system effectiveness.

4.6. Discussion on Dataset Readiness and Limitations.

The results show that gameplay logs from the first-grade mathematics educational game prototype can be transformed into analytical datasets through preprocessing and feature engineering. The validation results indicate that the raw data structure was sufficiently consistent for initial dataset construction. The cleaning and missing value interpretation results show that gameplay logs should be understood according to event context, rather than only by the presence or absence of values in specific fields. The feature engineering process then produced performance-related indicators and preliminary adaptive features that can support further analysis.

Dataset readiness can be seen from the formation of several analytical units, namely the answer-level dataset, level-level dataset, student-level dataset, and adaptive feature dataset. This multi-level structure enables gameplay logs to be analyzed from different perspectives. The answer-level dataset provides detailed information about responses to individual questions. The level-level dataset describes performance in each game mission. The student-level dataset summarizes overall gameplay performance. The adaptive feature dataset provides preliminary features that may be useful for future adaptive learning model development.

The main contribution of this article lies in dataset construction, not in testing game effectiveness. This article reinforces the position of gameplay logs as learning activity data that can be processed into structured datasets. In adaptive mathematics learning research, clean and structured datasets are required before adaptive models can be designed or evaluated. Without preprocessing and feature engineering, gameplay logs remain technical activity records that are difficult to use for learning analytics or data-driven model development.

This dataset has several limitations. First, the data were collected from a limited trial involving nine students, so the findings cannot be generalized to a broader population. Second, the dataset was obtained from one prototype and one learning context, so the variation of data remains limited. Third, the relatively high student accuracy suggests that the difficulty variation should be expanded in future game development. Fourth, the student_state and adaptation_action variables were still constructed using an initial rule-based approach, so they do not represent a final adaptive model. Therefore, this dataset should be positioned as an initial gameplay log dataset for validating data structure and constructing analytical features, not as evidence of learning effectiveness or final adaptive model performance.

5. CONCLUSIONS

This study constructed an initial gameplay log dataset from a first-grade elementary mathematics educational game prototype through a structured preprocessing and feature engineering procedure. The main focus of this study was to transform raw gameplay activity records into clean, organized, and analyzable datasets that can support early learning data analysis and provide an initial basis for future

Vol. 1 No. 4 May 2026
E-ISSN: 3132-5115
DOI: 10.65780/bima.v1i4.23

adaptive mathematics learning development. The dataset was obtained from a limited trial involving nine first-grade students who completed six sequential game levels.

The validation results showed that the raw gameplay log consisted of 262 records, 20 initial columns, nine unique students, nine unique sessions, six recorded levels, and no duplicate records. The event distribution included 171 answer events, 28 help events, 54 level completion events, and nine game completion events. These findings indicate that the logging structure of the prototype was able to consistently capture essential gameplay activities during the limited trial.

The preprocessing stage produced a validated, cleaned, anonymized, and event-filtered gameplay log. Missing values were interpreted according to event context, meaning that empty values in the selected answer field for help, level completion, and game completion events were not treated as data errors. This context-aware interpretation is important because gameplay logs are event-based datasets in which each event type has a different analytical meaning.

The feature engineering process generated several analytical features, including level accuracy, average response time, total help usage, average attempt count, overall accuracy, performance category, student state, and adaptation action. The resulting outputs consisted of answer-level, level-level, student-level, and adaptive feature datasets. These outputs show that raw gameplay logs can be transformed into structured analytical datasets that provide multiple levels of analysis for early student performance profiling.

The descriptive findings should be interpreted carefully. Although all students completed the six game levels and the dataset recorded student answers, help usage, response time, and level completion, these results do not represent evidence of learning effectiveness. Similarly, student state and adaptation action should be understood as preliminary rule-based features, not as the final output of an artificial intelligence or Q-Learning model.

This study has several limitations. The dataset was derived from a small-scale trial involving only nine students, so the findings cannot be generalized to a broader population. The data were also obtained from a single prototype and a single learning context. In addition, the adaptive features were still constructed using an initial rule-based approach. Future studies should involve a larger dataset, enrich the variation of difficulty levels, and further develop adaptive learning models that can be evaluated more comprehensively.

Overall, this study contributes to the early stage of adaptive mathematics learning development by presenting a preprocessing and feature engineering workflow for transforming raw gameplay logs into structured analytical datasets. The main contribution of this article lies in the dataset construction process, which prepares gameplay activity data as an initial data foundation for future adaptive learning model development.

The anonymized dataset and data dictionary are made available to support reproducibility and reuse in future research.

DATA AVAILABILITY

The anonymized gameplay log dataset and supplementary materials are available in Zenodo at https://zenodo.org/records/20576432. The repository includes the cleaned gameplay log, answer-level dataset, level-level dataset, student-level dataset, adaptive feature dataset, data dictionary, ethical note, license, and citation metadata. All student identifiers were anonymized before publication, and no personally identifiable information is included.

REFERENCES

[1] N. N. Dan, L. T. B. T. Trung, N. T. Nga, and T. M. Dung, “Digital game-based learning in mathematics education at primary school level: A systematic literature review,” *Eurasia Journal of Mathematics, Science and Technology Education*, vol. 20, no. 4, p. em2423, Apr. 2024, doi: 10.29333/ejmste/14377.

[2] E. Debrenti, “Game-Based Learning experiences in primary mathematics education,” *Front. Educ. (Lausanne).*, vol. 9, Mar. 2024, doi: 10.3389/feduc.2024.1331312.

[3] Y. Gui, Z. Cai, Y. Yang, L. Kong, X. Fan, and R. H. Tai, “Effectiveness of digital educational game and game design in STEM learning: a meta-analytic review,” *Int. J. STEM Educ.*, vol. 10, no. 1, p. 36, May 2023, doi: 10.1186/s40594-023-00424-9.

[4] R. Hidayat *et al.*, “Online game-based learning in mathematics education among Generation Z: A systematic review,” *International Electronic Journal of Mathematics Education*, vol. 19, no. 1, p. em0763, Jan. 2024, doi: 10.29333/iejme/14024.

[5] M. F. Rivaldi and Y. I. Kurniawan, “Game Edukasi Pengenalan dan Pembelajaran Berhitung untuk Siswa Kelas 1 Sekolah Dasar,” *Jurnal Manajemen Informatika (JAMIKA)*, vol. 11, no. 1, pp. 47–59, Mar. 2021, doi: 10.34010/jamika.v11i1.4354.

[6] R. Windawati and H. D. Koeswanti, “Pengembangan Game Edukasi Berbasis Android untuk Meningkatkan hassil Belajar Siswa di Sekolah Dasar,” *Jurnal Basicedu*, vol. 5, no. 2, pp. 1027–1038, Mar. 2021, doi: 10.31004/basicedu.v5i2.835.

[7] S. Damarjati and A. Miatun, “Pengembangan Game Edukasi Berbasis Android sebagai Media Pembelajaran Berorientasi pada Kemampuan Berpikir Kritis,” *ANARGYA: Jurnal Ilmiah Pendidikan Matematika*, vol. 4, no. 2, Oct. 2021, doi: 10.24176/anargya.v4i2.6442.

[8] S. K. Banihashem, H. Dehghanzadeh, D. Clark, O. Noroozi, and H. J. A. Biemans, “Learning analytics for online game-Based learning: a systematic literature review,” *Behaviour & Information Technology*, vol. 43, no. 12, pp. 2689–2716, Sep. 2024, doi: 10.1080/0144929X.2023.2255301.

[9] W. Lu, J. Griffin, T. D. Sadler, J. Laffey, and S. P. Goggins, “Game-Based Learning Prediction Model Construction,” *Journal of Learning Analytics*, vol. 12, no. 1, pp. 293–321, Mar. 2025, doi: 10.18608/jla.2025.8105.

[10] V. Shute *et al.*, “Maximizing learning without sacrificing the fun: Stealth assessment, adaptivity and learning supports in educational games,” *J. Comput. Assist. Learn.*, vol. 37, no. 1, pp. 127–141, Feb. 2021, doi: 10.1111/jcal.12473.

[11] C.-H. Chen, V. Law, and K. Huang, “Adaptive scaffolding and engagement in digital game-based learning,” *Educational technology research and development*, vol. 71, no. 4, pp. 1785–1798, Aug. 2023, doi: 10.1007/s11423-023-10244-x.

[12] R. Krisdiawan, T. Sugiharto, N. AmaliaAsikin, and L. Rohmawati, “UI/UX Design for AR Card Game: Enhancing English Vocabulary Learning with Augmented Reality,” *Kinetik: Game*

Vol. 1 No. 4 May 2026
E-ISSN: 3132-5115
DOI: 10.65780/bima.v1i4.23

Technology, Information System, Computer Network, Computing, Electronics, and Control, Jun. 2025, doi: 10.22219/kinetik.v10i3.2240.

[13] M. S. Alotaibi, “Game-based learning in early childhood education: a systematic review and meta-analysis,” *Front. Psychol.*, vol. 15, Apr. 2024, doi: 10.3389/fpsyg.2024.1307881.

[14] Y. Pan, F. Ke, and X. Xu, “A systematic review of the role of learning games in fostering mathematics education in K-12 settings,” *Educ. Res. Rev.*, vol. 36, p. 100448, Jun. 2022, doi: 10.1016/j.edurev.2022.100448.

[15] I. Daoudi, “Learning analytics for enhancing the usability of serious games in formal education: A systematic literature review and research agenda,” *Educ. Inf. Technol. (Dordr).*, vol. 27, no. 8, pp. 11237–11266, Sep. 2022, doi: 10.1007/s10639-022-11087-4.

[16] A. Tlili *et al.*, “A Systematic Literature Review of Empirical Studies on Learning Analytics in Educational Games.,” *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 7, no. 2, pp. 250–261, Dec. 2021, doi: 10.9781/ijimai.2021.03.003.

[17] T. J. E. Faber, M. E. W. Dankbaar, W. W. van den Broek, L. J. Bruinink, M. Hogeveen, and J. J. G. van Merriënboer, “Effects of adaptive scaffolding on performance, cognitive load and engagement in game-based learning: a randomized controlled trial,” *BMC Med. Educ.*, vol. 24, no. 1, p. 943, Aug. 2024, doi: 10.1186/s12909-024-05698-3.

[18] D. Chiotaki, V. Poulopoulos, and K. Karpouzis, “Adaptive game-based learning in education: a systematic review,” 2023, *Frontiers Media SA*. doi: 10.3389/fcomp.2023.1062350.

[19] W. Lu, J. Griffin, T. D. Sadler, J. Laffey, and S. P. Goggins, “Serious Game Analytics by Design: Feature Generation and Selection Using Game Telemetry and Game Metrics,” *Journal of Learning Analytics*, vol. 10, no. 1, pp. 168–188, Mar. 2023, doi: 10.18608/jla.2023.7681.

[20] A. Calvo-Morata, C. Alonso-Fernández, J. Santilario-Berthilier, I. Martínez-Ortiz, and B. Fernández-Manjón, “Learning Analytics to Guide Serious Game Development: A Case Study Using Articoding,” *Computers*, vol. 14, no. 4, p. 122, Mar. 2025, doi: 10.3390/computers14040122.