JOURNAL OF INFORMATION TECHNOLOGY AND ITS UTILIZATION, VOLUME 8, ISSUE 1, JUNE 2025
EISSN 2654-802X ; PISSN 2985-4067
DOI: https://doi.org/10.56873/jitu.8.1.5906. SUBMITTED: DECEMBER 25, 2024; REVISED: MARCH 6, 2025; ACCEPTED: MARCH 14, 2025

# Learning Difficulty Levels Prediction of Elementary School Student Mathematics Using Machine Learning Model

Rismayani[1*], Novita Sambo Layuk[2], Madyana Patasik[3], Andi Hutami Endang[4]

[1]Information System Study Program, Universitas Dipa Makassar, Indonesia
[2] Informatics Management Study Program, Universitas Dipa Makassar, Indonesia
[3]Informatics Study Program, Universitas Dipa Makassar, Indonesia
[4]Information System Study Program, Institut Teknologi dan Bisnis Kalla, Indonesia
rismayani@undipa.ac.id
*Corresponding author

***Abstract*--Difficulty learning mathematics in elementary school students is a significant problem and requires serious attention. This study aims to predict the difficulty level in elementary school students learning mathematics using a machine learning model, namely KNN. Exam scores, assignments, quizzes, and characteristics of students' difficulty level in learning mathematics were used as data in this study. A study used the KNN model to divide students into three categories of difficulty in learning mathematics: easy, moderate, and challenging. The results showed that the KNN model can accurately predict student's difficulty levels in mathematics. Thus, applying this model can help teachers provide appropriate and effective interventions to students experiencing difficulties. Using machine learning technology, especially the KNN model, we found an accuracy of 95%. In addition, we can still accurately predict the difficulty level of elementary school students' mathematics learning. This study uses anonymous student data, the distribution of assignments, quizzes, and exam score ranges, and characteristics of mathematics learning difficulty levels. There are three prediction classes: high, medium, and low.**

***Key words*: Mathematics learning; KNN (K-Nearest Neighbors) model; Difficulty level prediction; Educational intervention.**

## I. INTRODUCTION

As an important subject in education, mathematics fosters logical, analytical, and problem-solving abilities [1]. Mathematics is an essential basic science and a tool for other sciences, but it remains autonomous and does not depend on other fields of study [2] and [3]. According to the Indonesian Minister of Education and Culture Regulation No. 22 of 2006, mathematics is taught to all students to develop logical, analytical, systematic, critical, creative, and collaborative thinking skills [4]. However, many students, especially elementary school students, still struggle to learn mathematics. This can affect students' academic development and learning motivation. Internal factors that can cause difficulties in learning mathematics include cognitive ability, learning style, learning motivation, and interest in mathematics. External factors can also include the quality of teaching, learning environment, and family support. Conventional methods, such as value analysis, teacher observation, and diagnostic tests, have weaknesses due to subjectivity, time constraints, and lack of comprehensive data.

With the advancement in Machine Learning technology, there are more objective, effective, and comprehensive solutions to find problems in learning mathematics [5]. Machine Learning models can make accurate predictions and find hidden patterns in extensive, complex data. This study uses the KNN (K-Nearest Neighbors) machine learning model to determine how difficult it is for elementary school students to learn mathematics. The KNN model was chosen because it can classify data based on similarities with previously known data.

The research problem formulation is how to evaluate the ability of a machine learning model such as KNN to predict the difficulty level in learning mathematics for elementary school students. Thus, the study aims to predict the difficulty level in elementary school students learning mathematics using a machine learning model, namely KNN, by analyzing and evaluating accuracy.

Predicting elementary school students' mathematics learning difficulties using Machine Learning, specifically the K-Nearest Neighbors (KNN) model, has shown promising results. This approach utilizes multiple predictors to classify

EISSN 2654-802X ; PISSN 2985-4067
DOI: https://doi.org/10.56873/jitu.8.1.5906. SUBMITTED: DECEMBER 25, 2024; REVISED: MARCH 6, 2025; ACCEPTED: MARCH 14, 2025

students' performance levels effectively. Several previous studies conducted by the author on KNN can classify students based on performance metrics, such as test scores and engagement levels, allowing educators to predict students at risk of underachieving [6]. Qualitative studies in which students' difficulties were categorized based on their problem-solving approaches showed that KNN can effectively classify students into different performance levels by looking at how they respond to mathematical problems [7]. Next, research that analyzes reliability learning strategies to improve student motivation and learning outcomes [8]. This research compared different machine learning techniques to predict student graduation. This machine learning method starts with data collection, transformation, and analysis [9]. The research found that if the teacher's teaching method is suitable, the grade is B, the student's interest in learning is low, and the gender is male. The student's level of understanding in mathematics lessons is low [10]. Then, there is also research that utilizes machine learning technology, such as research that predicts plant seeds [11] which provides useful plant information for urban agriculture management. Research that predicts harmful content from Twitter [12] where the models used are TF-IDF, Logistics Regression Model, and Naive Bayes Classifier for sentiment classification. Furthermore, research in other fields that apply machine learning to predictions, such as research discussing predictions of nervous disorders, is also available [13] and on eye disorders [14]. According to the research, the WebQual 4.0 method is used to evaluate the implementation of educational website-based information system services. This method evaluates user satisfaction based on usability, information quality, and service interaction quality [15]. Thus, recommendations for development and improvement can be made to evaluate the implementation of educational website-based information systems.

In the previous research conducted by the author, several discussions of mathematics learning were followed by the application of machine learning, KNN, and the prediction process. Still, the prediction of the level of learning difficulty for elementary school students has not been studied much, especially in evaluating the performance of the model used.

This study significantly improves the quality of mathematics education in elementary schools. First, this study provides a better understanding of the problems faced by elementary school students when learning mathematics. This study predicts patterns and factors contributing to learning difficulties that may be missed if only traditional methods are used. Second, this study helps teachers find students who need additional help and provide appropriate interventions. Teachers can create more unique and effective learning strategies by knowing the level of difficulty faced by each of their students. These three studies show how machine learning technology can improve the quality of mathematics education. Machine learning models can analyze large and complex data, find hidden patterns, and provide teachers with important information. Finally, it improves accuracy, predicts learning problems better, and helps teachers choose the most appropriate action for the problem.

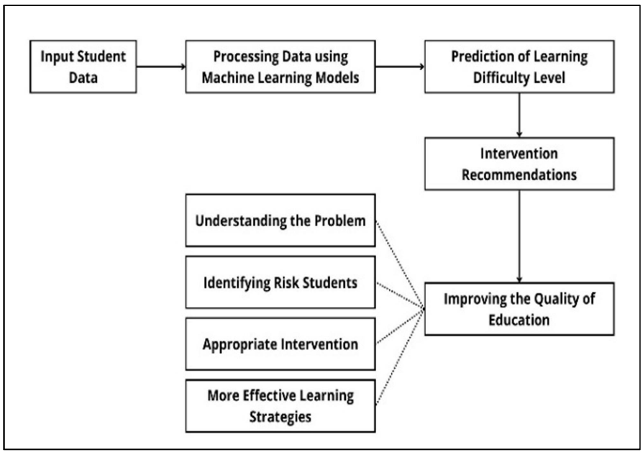


Fig. 1. Framework Diagram

Fig.1 uses a conceptual framework that combines student data with a machine learning model to predict the level of mathematics learning difficulties. Student data, including test scores, ability test results, demographic data, and family information, are processed using a specially designed machine learning algorithm to identify patterns and factors contributing to learning difficulties. The model generates predictions about each student's level of learning difficulties, both in categories (low, medium, high) and numerical values. Based on these predictions, the

EISSN 2654-802X ; PISSN 2985-4067
DOI: https://doi.org/10.56873/jitu.8.1.5906. SUBMITTED: DECEMBER 25, 2024; REVISED: MARCH 6, 2025; ACCEPTED: MARCH 14, 2025

system recommends appropriate interventions, such as additional learning programs, individual tutoring, or customized teaching strategies. Thus, this study aims to improve the quality of mathematics education by understanding students' problems, identifying at-risk students, providing targeted interventions and helping teachers create more effective learning strategies.

## II. METHOD

At this stage, the methods used in the research are explained, starting from the scenario model used, data collection, research tools and materials, models, and evaluation.

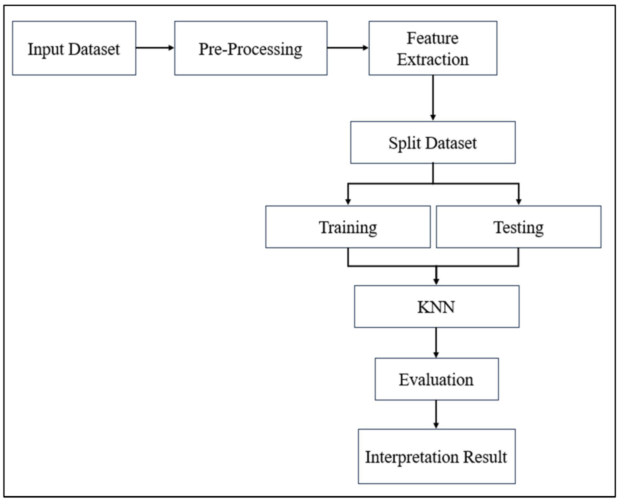


Fig. 2. Research Methodology

Fig. 2 is a scenario of the research methodology. The first process is to input the dataset. Then, the dataset is pre-processed, after which the features are extracted. The features in question are assignment scores, quiz scores, exam scores, and characteristics of difficulty level in learning mathematics. Furthermore, the dataset is split, namely 80% training and 20% testing. The next stage is modeling with KNN, which is evaluated to determine its accuracy and display the results.

### A. Dataset Collection

The dataset consists of 200 data obtained from elementary school student respondents from grades 4 to 6 after undergoing a mathematics learning process. The data collected were in the form of assignment scores, quiz scores, exam scores, and respondent answers regarding the characteristics of the difficulty level of mathematics learning. The dataset obtained was split into 80% training and 20% testing. Table I shows an example of the dataset used as follows:

TABLE I
Samples of Dataset

| Student | Assignment Value | Quiz Score | Final Exam Score | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 | C9 | C10 | C11 | C12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| S1 | 87 | 85 | 80 | Yes | Yes | Yes | Yes | Yes | Yes | Yes | No | Yes | No | Yes | No |
| S2 | 70 | 75 | 80 | No | No | Yes | Yes | No | No | Yes | Yes | Yes | No | Yes | No |
| S3 | 80 | 70 | 71 | Yes | Yes | No | No | No | Yes | No | Yes | Yes | Yes | No | Yes |
| S4 | 50 | 40 | 45 | No | Yes | Yes | Yes | Yes | No | Yes | Yes | Yes | Yes | No | No |
| S5 | 60 | 70 | 70 | No | No | No | No | No | Yes | Yes | Yes | No | Yes | Yes | No |
| S6 | 90 | 85 | 87 | Yes | Yes | No | No | Yes | Yes | No | No | Yes | No | No | Yes |
| S7 | 85 | 85 | 85 | Yes | No | Yes | Yes | No | No | Yes | Yes | Yes | Yes | No | No |
| S8 | 50 | 50 | 50 | Yes | No | Yes | Yes | Yes | Yes | No | No | Yes | Yes | Yes | Yes |
| S9 | 75 | 70 | 65 | No | Yes | No | No | Yes | Yes | Yes | Yes | No | No | No | No |
| S10 | 80 | 83 | 85 | No | No | Yes | Yes | No | No | No | No | Yes | Yes | Yes | Yes |
| S11 | 87 | 85 | 80 | Yes | Yes | Yes | Yes | Yes | Yes | Yes | No | Yes | No | Yes | No |
| S12 | 70 | 75 | 80 | No | No | Yes | Yes | No | No | Yes | Yes | Yes | No | Yes | No |
| S13 | 80 | 70 | 71 | Yes | Yes | No | No | No | Yes | No | Yes | Yes | Yes | No | Yes |
| S14 | 50 | 40 | 45 | No | Yes | Yes | Yes | Yes | No | No | No | Yes | No | Yes | Yes |
| S15 | 60 | 70 | 70 | No | No | No | No | No | Yes | Yes | Yes | No | Yes | Yes | No |
| … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … |
| S200 | 50 | 75 | 75 | No | No | No | No | No | Yes | Yes | Yes | No | Yes | Yes | No |

Information:
S1 – S200 = Student; C1 – C12 = Characteristic features

EISSN 2654-802X ; PISSN 2985-4067
DOI: https://doi.org/10.56873/jitu.8.1.5906. SUBMITTED: DECEMBER 25, 2024; REVISED: MARCH 6, 2025; ACCEPTED: MARCH 14, 2025

### *B. Pre-Processing*

The collected dataset is pre-processed before the feature extraction using the data cleaning and data normalization method; then, the data is extracted as in Table II, III and IV.

The research tools and materials used to support the research are software and hardware tools, including Google Collab, Python 3 Programming, Adam Optimizer, and a learning rate 0.0001. The hardware is an AMD Ryzen 7 5825U, with 16GB memory and 500GB HDD. The research materials used are primary data, such as elementary school student assignment data, quiz data, exam data, and data on students' learning difficulties.

TABLE II
Range of Value Feature Distribution

| Num | Assignment Value | Quiz Score | Final Exam Score | Value |
|---|---|---|---|---|
| 1 | 85-100 | 85-100 | 85-100 | 0 |
| 2 | 50 - 84 | 50 - 84 | 50 - 84 | 1 |
| 3 | < 50 | < 50 | < 50 | 2 |

TABLE III
Characteristic Feature

| Variable | Characteristic | Value |
|---|---|---|
| C1 | Difficulty calculating basic mathematical operations such as addition, subtraction, multiplication, and division. | Yes = 1<br>No = 0 |
| C2 | Difficulty understanding basic concepts such as whole numbers, fractions, or percentages. | Yes = 1<br>No = 0 |
| C3 | Difficulty in applying mathematical concepts in everyday life situations. | Yes = 1<br>No = 0 |
| C4 | Difficulty in remembering basic math formulas. | Yes = 1<br>No = 0 |
| C5 | Difficulty in understanding number patterns and sequences. | Yes = 1<br>No = 0 |
| C6 | Difficulty understanding more complex mathematical concepts such as algebra, geometry, or statistics. | Yes = 1<br>No = 0 |
| C7 | Difficulty in mastering the mathematical techniques and strategies needed to solve problems. | Yes = 1<br>No = 0 |
| C8 | Difficulty in understanding more complex mathematical notation. | Yes = 1<br>No = 0 |
| C9 | Difficulty understanding highly abstract mathematical concepts such as calculus, number theory, or linear algebra. | Yes = 1<br>No = 0 |
| C10 | Difficulty applying mathematical concepts in very complex situations or solving complicated mathematical problems. | Yes = 1<br>No = 0 |
| C11 | Difficulty in understanding more complex mathematical proofs and mathematical logic. | Yes = 1<br>No = 0 |
| C12 | Difficulty in remembering very complex mathematical formulas. | Yes = 1<br>No = 0 |

TABLE IV
Normalization Data

| Student | Assignment Value | Quiz Score | Final Exam Score | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 | C9 | C10 | C11 | C12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| S1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 |
| S2 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 0 |
| S3 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 1 |
| S4 | 1 | 2 | 2 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 |
| S5 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 |
| S6 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 |
| S7 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 |
| S8 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 |
| S9 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |

EISSN 2654-802X ; PISSN 2985-4067
DOI: https://doi.org/10.56873/jitu.8.1.5906. SUBMITTED: DECEMBER 25, 2024; REVISED: MARCH 6, 2025; ACCEPTED: MARCH 14, 2025

| Student | Assignment Value | Quiz Score | Final Exam Score | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 | C9 | C10 | C11 | C12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| S10 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 |
| S11 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 |
| S12 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 0 |
| S13 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 1 |
| S14 | 1 | 2 | 2 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 |
| S15 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 |
| … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … |
| S200 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 |

### C. *Split Dataset*

At this stage, the dataset is separated into training and testing data, with the proportions of each data shown in Table V and Fig.3.

TABLE V
Split Data

| Split Data | Amount | Percentage |
|---|---|---|
| Training | 160 | 80% |
| Testing | 40 | 20% |
| Total | 200 | |

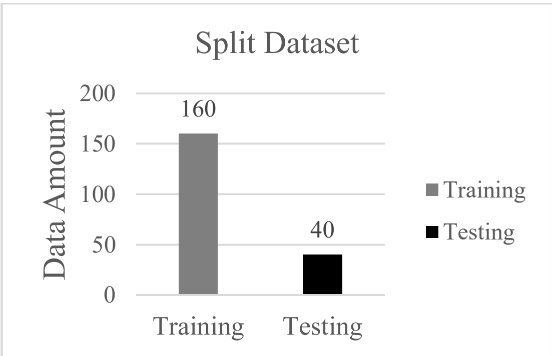


Fig. 3. Split Dataset

### D. *Model Training and Testing*

This stage explains the model or technology used to predict the difficulty level of learning mathematics. The technology used is machine learning. The model used is KNN, a simple but effective Machine Learning algorithm for predicting values or classifying data [16]. This algorithm finds the nearest K data points (neighbors) of a new data point to be classified or predicted. The new data point is then classified into the majority class of the K nearest neighbors, or the average value of the K nearest neighbors is used as a prediction. The working steps of the KNN model are as follows:

1. Determine the value of K, the number of nearest neighbors considered in the prediction process.
2. Calculate the distance between points with Euclidean Distance with the formula equal (1).

$$Distance = \sqrt{((x1 - x2)^2 + (y1 - y2)^2)} \qquad (1)$$

Where:

- (x1, y1): coordinates of new data point
- (x2, y2): coordinates of data point in the training set

3. Select the nearest K based on the calculated distance.
4. Determine the class that is most of the K nearest neighbors.

### E. *Evaluation*

This stage evaluates the model used by calculating the accuracy matrix with an accuracy formula such as equation (2).

$$Accuracy = \frac{(Number\ of\ Correct\ Predictions)}{(Total\ Number\ of\ Predictions)} \times 100 \qquad (2)$$

Where:

- Number of correct predictions: Number of data points correctly classified or predicted by the KNN model.
- Total Number of predictions: Total Number of data points predicted by the KNN model.

### F. *Interpretation Result*

Table VI shows the interpretation stage of the results, which is the predicted level of difficulty in learning mathematics.

JOURNAL OF INFORMATION TECHNOLOGY AND ITS UTILIZATION, VOLUME 8, ISSUE 1, JUNE 2025
EISSN 2654-802X ; PISSN 2985-4067
DOI: https://doi.org/10.56873/jitu.8.1.5906. SUBMITTED: DECEMBER 25, 2024; REVISED: MARCH 6, 2025; ACCEPTED: MARCH 14, 2025

TABLE VI
Mathematics Learning Difficulty Level Prediction Class

| variables | Value | Target |
|---|---|---|
| K1 | 0 | High |
| K2 | 1 | Medium |
| K3 | 2 | Low |

## III. RESULT AND DISCUSSION

At this stage, the results and discussion of the research conducted. The results of the prediction experiment are displayed based on the training data that has been carried out with the KNN model in Table VII, where the value results of each variable are by the experimental results that have been processed as follows:

TABLE VII
Experiment Prediction

| Student | Assignment Value | Quiz Score | Final Exam Score | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 | C9 | C10 | C11 | C12 | Result |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| S1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | K1 |
| S2 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | K2 |
| S3 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | K2 |
| S4 | 1 | 2 | 2 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | K3 |
| S5 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | K2 |
| S6 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | K1 |
| S7 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | K1 |
| S8 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | K2 |
| S9 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | K2 |
| S10 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | K2 |
| S11 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | K1 |
| S12 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | K2 |
| S13 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | K2 |
| S14 | 1 | 2 | 2 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | K3 |
| S15 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | K2 |
| … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … |
| S200 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | K2 |

Table VII shows the prediction results of the experiment. The data consists of 200 students (S1 to S200). For each student, there is data on assignment value, quiz score, and final exam score. Columns C1 to C12 represent 12 variables or features used in the prediction model (details of the features are not explained in the table). The values 0 and 1 in columns C1-C12 may indicate the presence or absence of a student characteristic or attribute relevant to mathematics learning difficulty. The "Result" column shows the classification of the prediction results for each student, categorized as K1, K2, or K3 (possibly representing different levels of learning difficulty, such as low, medium, and high). The "..." sign indicates that the data for students between S15 and S200 are similar to the pattern shown above. This table provides empirical data to evaluate the machine learning prediction model's performance in classifying students' mathematics learning difficulty levels.

TABLE VIII
Experiment

| Student | AS | QS | FES | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 | C9 | C10 | C11 | C12 | Result | P |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| S1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | K1 | K1 |
| S2 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | K2 | K2 |
| S3 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | K2 | K2 |
| S4 | 1 | 2 | 2 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | K3 | K3 |
| S5 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | K2 | K2 |
| S6 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | K1 | K2 |
| S7 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | K1 | K1 |
| S8 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | K2 | K2 |

EISSN 2654-802X ; PISSN 2985-4067
DOI: https://doi.org/10.56873/jitu.8.1.5906. SUBMITTED: DECEMBER 25, 2024; REVISED: MARCH 6, 2025; ACCEPTED: MARCH 14, 2025

| Student | AS | QS | FES | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 | C9 | C10 | C11 | C12 | Result | P |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| S9 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | K2 | K2 |
| S10 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | K2 | K1 |
| S11 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | K1 | K1 |
| S12 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | K2 | K2 |
| S13 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | K2 | K2 |
| S14 | 1 | 2 | 2 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | K3 | K3 |
| S15 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | K2 | K2 |
| … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … |
| S200 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | K2 | K2 |

Description:
AS: Assignment Value
QS: Quiz Score
FES: Final Exam Score
P: Prediction

Table VIII shows the prediction results of an experiment aimed at predicting the level of mathematics learning difficulty in 200 students (S1-S200), continuing the same data pattern up to S200. The input data include assignment (AS), quiz (QS), and final exam (FES) scores, which may be on a binary scale (0/1), and 12 variables (C1-C12) representing student characteristics, also on a binary scale, but the details are not explained. The prediction model produces two classifications of learning difficulty: "Result" (K1, K2, K3, possibly representing low, medium, and high levels) and "P," which may represent predictions from a different model or revised predictions. The differences between "Result" and "P," especially those highlighted in red, indicate mismatched predictions and highlight the potential complexity of predicting mathematics learning difficulty.

TABLE IX
Prediction Validation

| Student | AS | QS | FES | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 | C9 | C10 | C11 | C12 | Result | P | Inf. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| S1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | K1 | K1 | V |
| S2 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | K2 | K2 | V |
| S3 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | K2 | K2 | V |
| S4 | 1 | 2 | 2 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | K3 | K3 | V |
| S5 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | K2 | K2 | V |
| S6 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | K1 | K2 | In |
| S7 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | K1 | K1 | V |
| S8 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | K2 | K2 | V |
| S9 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | K2 | K2 | V |
| S10 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | K2 | K1 | In |
| S11 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | K1 | K1 | V |
| S12 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | K2 | K2 | V |
| S13 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | K2 | K2 | V |
| S14 | 1 | 2 | 2 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | K3 | K3 | V |
| S15 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | K2 | K2 | V |
| … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … |
| S40 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | K2 | K2 | V |

Description: Inf.: Information; V: Valid; In: Invalid

Table IX presents the prediction validation results of an experiment predicting mathematics learning difficulties in 40 students (S1-S40), with a continuous data pattern up to S40. Similar to Table VIII, the input data include assignment scores (AS), quizzes (QS), final exam scores (FES), and 12 variables (C1-C12) representing student characteristics, all using a binary scale (0/1). The prediction model produces two classifications of learning difficulties: "Result" (K1, K2, K3) and "P," which may represent predictions from different models or revised predictions. The "Inf" column likely provides additional information about the validity of the

EISSN 2654-802X ; PISSN 2985-4067
DOI: https://doi.org/10.56873/jitu.8.1.5906. SUBMITTED: DECEMBER 25, 2024; REVISED: MARCH 6, 2025; ACCEPTED: MARCH 14, 2025

predictions, with "V" and "In" possibly representing validation and invalidation of predictions, respectively. Differences between "Result" and "P," particularly those indicated by different colors, indicate prediction discrepancies

$$Accuracy = \frac{38}{40} \times 100 = 95\%.$$

So that, the accuracy obtained is a rate of 0.95 or 95%. Other authors have conducted several comparative studies regarding predictions that apply machine learning models that discuss using the K-Means model with KNN in identifying student knowledge [17] The data still uses data from one campus only in one department and does not evaluate the accuracy of the model used, so the model's accuracy level for the research topic is not yet visible. Next, the research discusses predictions of user experience regarding the Melisa application using SVM and KNN [18], the analysis is carried out to determine the user experience, focusing on digital products and improving the UI/UX appearance and content. Research that discusses the prediction of elementary school students' achievements using C.45 [19] where accuracy has not been explained as to what is produced. Then, the research predicts prospective student performances during the entrance selection process [20] which produces quite good accuracy for various Machine Learning models.

The study implications show that compared to the traditional approach, the KNN Machine Learning model used in this study can improve the accuracy of predicting students' mathematics learning difficulties. The results of this study can help teachers find students who have difficulty learning mathematics and offer appropriate solutions. This study also opens up new opportunities for using Machine Learning technology in education, improving the quality of learning, increasing teaching effectiveness, and helping teachers understand and overcome students' learning difficulties more accurately by predicting learning difficulties.

The study limitations are that the research dataset used is still lacking because it takes samples from several elementary schools in the Maros district, and the data quality still contains bias. Furthermore, KNN is a model that is still and highlight the complexity of predicting mathematics learning difficulties. This table allows for the evaluation of the prediction model's performance and analysis of the validity of the predictions.
simple, and external factors can still affect the difficulty of learning mathematics.

## IV. Conclusion

This study successfully demonstrated that machine learning models, especially K-Nearest Neighbors (KNN), can be used effectively to predict the level of learning difficulty in mathematics in elementary school students. By analyzing data obtained from test scores, assignments, quizzes, and student characteristics, the KNN model could classify students into three categories of difficulty levels: easy, medium, and complex, with an accuracy of 95%. These results confirm the potential use of machine learning technology in education, especially in helping teachers provide appropriate and effective interventions for students who experience learning difficulties. Thus, applying this model can improve students' understanding of mathematics and support their overall academic development. This study also emphasizes the importance of collecting anonymous and relevant data to improve prediction accuracy in the educational context.

Future work is to collect more complete data on a broader scope. Conduct testing with various models from relevant Machine learning and consider variables from external factors.

## V. Acknowledgment

The author would like to thank DRPM DIKSI with the main contract number 106/SPK/D.D4/PPK.01.APTV/III/2024 and the derivative contract 259/LL9/PK.00.PGPV/2024, which has provided funding for this research so that this research can be completed properly.

## VI. References

[1] Ernawati *et al.*, *Problematika Pembelajaran Matematika*. Yayasan Penerbit Muhammad Zaini, 2021.

[2] Isrok'atun and A. Rosmala, *Model-Model Pembelajaran Matematika*. Bumi Aksara, 2021.

[3] Z. Ardi *et al.*, "Exploring the elementary students learning difficulties risks on mathematics based on students mathematic anxiety, mathematics self-efficacy

EISSN 2654-802X ; PISSN 2985-4067
DOI: https://doi.org/10.56873/jitu.8.1.5906. SUBMITTED: DECEMBER 25, 2024; REVISED: MARCH 6, 2025; ACCEPTED: MARCH 14, 2025

and value beliefs using rasch measurement," *J. Phys. Conf. Ser.*, vol. 1157, no. 3, p. 032095, Feb. 2019, doi: 10.1088/1742-6596/1157/3/032095.

[4] M. Yusuf Setia Wardana and Aries Tika Damayani, "Persepsi Siswa terhadap Pembelajaran Pecahan di Sekolah Dasar," *ResearchGate*, Oct. 2024, doi: 10.31980/mosharafa.v6i3.333.

[5] S. Susanti *et al.*, *Desain Media Pembelajaran SD/MI*. Yayasan Penerbit Muhammad Zaini, 2022.

[6] J. L. Harvey and S. A. P. Kumar, "A Practical Model for Educators to Predict Student Performance in K-12 Education using Machine Learning," in *2019 IEEE Symposium Series on Computational Intelligence (SSCI)*, Dec. 2019, pp. 3004–3011. doi: 10.1109/SSCI44817.2019.9003147.

[7] G. S. Sidik and A. A. Wakih, "Kesulitan Belajar Matematik Siswa Sekolah Dasar pada Operasi Hitung Bilangan Bulat," *Nat. J. Kaji. Dan Penelit. Pendidik. Dan Pembelajaran*, vol. 4, no. 1, pp. 461–470, 2019, doi: 10.35568/naturalistic.v4i1.633.

[8] M. Anggriyani, M. Mahsup, S. Sirajuddin, and S. Syaharuddin, "Pembelajaran Estafet Learning dengan Kemampuan Numerik untuk Meningkatkan Motivasi dan Hasil Belajar Siswa," *Semin. Nas. Paedagoria*, vol. 4, no. 1, pp. 443–452, Aug. 2024.

[9] A. H. Anna Zili, "Prediksi Kelulusan Siswa pada Mata Pelajaran Matematika menggunakan Educational Data Mining," *J. Ris. Pembelajaran Mat. Sekol.*, vol. 7, no. 1, Mar. 2023, doi: 10.21009/jrpms.071.03.

[10] A. Patrisyah, R. Buaton, and J. N. Sitompul, "Klasifikasi Tingkat Pemahaman Siswa pada Pelajaran Matematika di MTSS PAB 5 Klambir Lima," *Saturnus J. Teknol. Dan Sist. Inf.*, vol. 2, no. 4, pp. 146–156, Aug. 2024, doi: 10.61132/saturnus.v2i4.345.

[11] C. Chazar and M. H. Rafsanjani, "Penerapan \Teachable Machine Pada Klasifikasi Machine Learning Untuk Identifikasi Bibit Tanaman," *Pros. Semin. Nas. Inov. Dan Adopsi Teknol. INOTEK*, vol. 2, no. 1, pp. 32–40, May 2022, doi: 10.35969/inotek.v2i1.207.

[12] A. K. Sari, A. Irsyad, D. N. Aini, Islamiyah, and S. E. Ginting, "Analisis Sentimen Twitter Menggunakan Machine Learning untuk Identifikasi Konten Negatif," *Adopsi Teknol. Dan Sist. Inf. ATASI*, vol. 3, no. 1, Art. no. 1, Jun. 2024, doi: 10.30872/atasi.v3i1.1373.

[13] Rismayani, S. R. D. Rachman, S. Wahyuni, Asmanurhidayani, J. Y. Mambu, and M. Pineng, "Implementation Artificial Neural Network on Identification System of Neurological Disorder," in *Intelligent Communication Technologies and Virtual Mobile Networks*, G. Rajakumar, K.-L. Du, and Á. Rocha, Eds., Singapore: Springer Nature, 2023, pp. 619–629. doi: 10.1007/978-981-99-1767-9_45.

[14] R. Rismayani, M. Pineng, and H. Herlinda, "Using Artificial Neural Network for System Education Eye Disease Recognition Web-Based," *J. Biomim. Biomater. Biomed. Eng.*, vol. 55, pp. 262–274, 2022, doi: 10.4028/p-7z9xpt.

[15] N. H. Sutanto, E. Utami, and R. Rismayani, "Systematic Literature Review untuk Identifikasi Metode Evaluasi Website Layanan Pendidikan di Indonesia," *J. Ilm. IT CIDA*, vol. 7, no. 1, Art. no. 1, Jun. 2021, doi: 10.55635/jic.v7i1.133.

[16] R. Aprilian, R. Habibi, and M. Y. H. Setyawan, *Algoritma KNN dalam memprediksi cuaca untuk menentukan tanaman yang cocok sesuai musim*. Kreatif, 2020.

[17] H. Sunandar, "Penggunaan Model Klaster K-Means Dan Klasifikasi KNN Untuk Identifikasi Pengetahuan Matematika Mahasiswa," *KAKIFIKOM Kumpul. Artik. Karya Ilm. Fak. Ilmu Komput.*, pp. 75–85, Apr. 2024.

[18] D. Kartikasari and I. K. D. Nuryana, "Analisis Prediksi Pengalaman Pengguna Aplikasi MELISA menggunakan Metode SVM dan KNN," *J. Emerg. Inf. Syst. Bus. Intell. JEISBI*, vol. 4, no. 4, pp. 72–78, Sep. 2023.

[19] T. Telutci and R. Harman, "Penerapan Data Mining untuk Memprediksi Prestasi Siswa Sekolah Dasar Menggunakan Algoritma C4.5," *Comput. Based Inf. Syst. J.*, vol. 12, no. 1, Art. no. 1, Mar. 2024, doi: 10.33884/cbis.v12i1.8207.

[20] S. D. Ariyanto and L. A. Wulandhari, "Prediksi Kinerja Calon Mahasiswa Berdasarkan Nilai Seleksi Masuk Menggunakan Pendekatan Machine Learning," *J. Ilm. Komputasi*, vol. 23, no. 2, Art. no. 2, Jun. 2024, doi: 10.32409/jikstik.23.2.3589.