Predicting Math Performance of Children with Special Needs Based on Serious Game

Umi Laili Yuhana[1,2], Remy G. Mangowal[1], Siti Rochimah[2], Eko M. Yuniarno[1], Mauridhi H. Purnomo[1]
[1]Department of Electrical Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
[2]Department of Informatics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
{yuhana16, remy.mangowal16}@mhs.ee.itc.ac.id, siti@if.its.ac.id, {ekomulyanto, hery}@ee.its.ac.id

Abstract — **Predicting and classifying student's performance using data mining techniques have been gaining an enormous amount of attention from researchers and practitioners. However, the use of games for the classification of student's ability level is still slightly. This study focuses on identification of important factors for determining student level performance on Math. The best classification algorithm is observed as part of intelligent game development research for assessment of children with special needs. The real dataset from randomly selected of elementary school is taken to construct a dataset. About 135 normal students and 25 children with special needs played the game and did a manual test. Our study shows that the age, gender, grade, and mark of each level became important factors in determining the level of math skill for the normal student. However age, gender, and grade don't have a correlation with math level of children with special needs. Six classification methods, Naive Bayes, Multilayer Perceptron (MLP), SMO, Decision Table, JRip, and J48, were performed to predict math skill performance level of normal students and children with special needs. JRip with 10 fold cross validation gives the highest percentage of accuracy of 64.12.**

Keywords—student performance prediction; math game; children with special needs.

I. INTRODUCTION

Predicting and classifying student's performance using data mining techniques have been gaining an enormous amount of attention from researchers and practitioners[1]. The outputs of classification that were researched are different. Some results indicated that the factors affecting student performance also varies. The classification and prediction results can be used as a consideration for determining the content of learning, assessment questions, as well as an early warning.

The approaches used to obtain the dataset also varies. The questionnaire is more widely used [2],[3],[4]. Some researchers used the mark of certain course and GPA as the basis of classification [3],[5]. However, the use of games to get the dataset is still slightly. Despite the negative effect of the gameplay, the presence of serious games can be used to help students get the benefit [6]. Through the game, the concept of learning while playing a game widely applied to improve the learning success. The rapid growth of gadgets and gadget utilization among children to play the game can be used to construct a dataset for classification based on log data of the game.

Mathematics is one of the basic lessons that must be understood by the students. But many students in Indonesia found the subject difficult to master. Statistical results of national examinations at elementary school and middle school in Indonesia showed that mathematics is one of the major causes of student failure in passing the national exam [7]. Mathematics became compulsory subject from elementary education, with the kind of competence that increases from level to level. Some concepts are interconnected between levels. For example, the concept of multiplication and division of numbers in grade 2, related to the concept of addition and subtraction that are given in grade 1.

Based on the opinion of the teachers, often found students have not mastered the concepts to their level. This will lead to a higher probability of the failure of students. Detection of precise classification of students' cognitive level on the Math will help teachers deliver appropriate content and methods for each student, especially for students with disabilities. Moreover, children with special needs on average have lower ability compared to normal children in the same age.

This paper focuses on the use of gameplay data for predicting math skill level of students. This study is part of intelligent game development research for assessment of children with special needs. The main objectives of this study are

- Identify important attributes that can be used to predict the student performance level in Math,
- find the best classification algorithm for predicting the student performance level.

II. LITERATURE STUDY

Ramesh et al have identified important factors that affect student performance in the final test and predicted students grade based on these factors [4]. Predicted results utilized to provide an appropriate warning for students who are at risk. Using 29 student-related variables, five data mining algorithm such as J48, SMO, REPTree, Naive Bayes and MLP were applied on dataset of 900 secondary students. It was reported that MLP algorithm had the best-predicted accuracy of 72.38%.

Harwati et al have been mapped student's performance to find a hidden pattern and classify the students based on their demographic data [3]. Six features were used as input for clustering. The profiles of 306 university students were collected as a dataset. Three clusters students (low, average, and smart student) were found using K-means clustering algorithm. This result could be used to improve the academic performance in Faculty of Industrial Engineering Department of industrial Technology, Islamic University of Indonesia.

Kaur et al used classification and prediction based data mining to identify slow learners [2]. Amount of 152 high school students in India were involved in observation that type learner. From 14 factors that may affect student performance, eight factors were identified as the most influential factors. The classification was done using five data mining techniques in WEKA, i.e. MLP, Naive Bayes, SMO, J48, and REPTree. The approaches successfully classify slow learners with the best accuracy, 75%, was obtained by using MLP.

The use of games for classification of student's ability level is still slightly. Sukajaya et al proposed bloom taxonomy based serious game to replace the paper based assessment [8]. 85 elementary students were involved in the study by playing the game. Classification of learner's cognitive skill has been done using 29 attributes from game log. Three data mining techniques namely Bayesian Network, Naive Bayes and J48 were performed for classification. This study found that Naive Bayes Classifier provided the best accuracy 92.31%.

Castellar et al observed the effectiveness of commercial educational math game for improving the arithmetic skill of children [9]. Seventy-four children in three gaming groups were observed. One group was instructed to play the game, one group was instructed to complete math exercise on paper and the last group did not receive any arithmetic exercises. They found that mental calculation speed can be improved by playing the educational math game such Monkey Tales [10].

III. METHODOLOGY

The methodology used in this study is shown in Fig 1. There are 3 data that used in this study; student's mark, written test result, and gameplay data. Gameplay data is collected using math game. This section discusses math game, data collection, data preprocessing, and scenario for classification and prediction.

A. Math Game

The dataset for prediction was constructed based on Math game, an assessment serious game which is developed to assess Math skill of children with special needs. This game adopts Indonesian Math curriculum for elementary school students. The game is presented in the form of a quiz in multiple choices with 4 options.

The data used for the game is a math question that consists of 60 questions. These questions are taken from final exam of grade 1 to grade 6 and have been validated by experts. Suppose that $q1$to $q60$ are question 1 to 60. As (1) Q is set of element of questions. Each 10 questions represent one grade level as in (2), (3), (4), (5), (6) and (7), where $QL1$ is set of questions in grade level 1, $QL2$ is set of questions in grade level 2, $QL3$ is set of questions in grade level 3, $QL4$ is set of questions in grade level 4, $QL5$ is set of questions in grade level 5, and $QL6$ is set of questions in grade level 6.

$$Q \in \{q1, q2, q3, q4, q5, q6, \ldots, q60\} \quad (1)$$

$$QL1 \in \{q1, q2, q3, q4, q5, q6, q7, q8, q9, q10\} \quad (2)$$

$$QL2 \in \{q11, q12, q13, q14, q15, q16, q17, q18, q19, q20\} \quad (3)$$

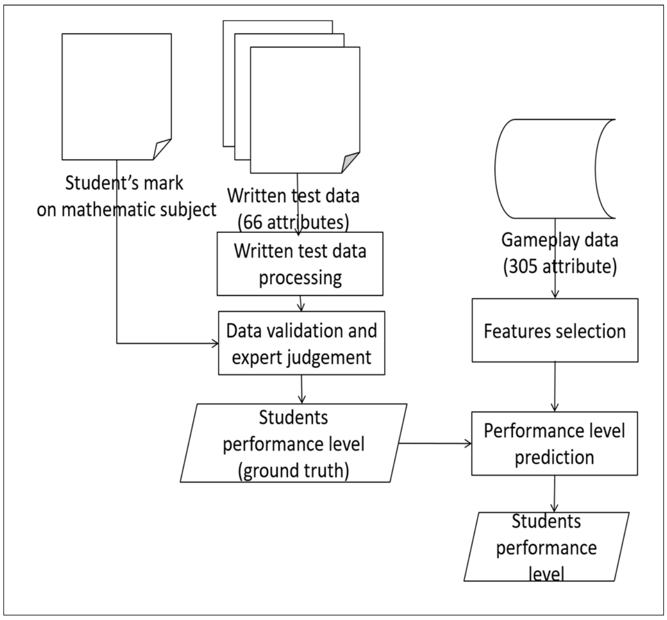

Fig. 1. Flowchart of proposed work

$$QL3 \in \{q21, q22, q23, q24, q25, q26, q27, q28, q29, q30\} \quad (4)$$

$$QL4 \in \{q31, q32, q33, q34, q35, q36, q37, q38, q39, q40\} \quad (5)$$

$$QL5 \in \{q41, q42, q43, q44, q45, q46, q47, q48, q49, q50\} \quad (6)$$

$$QL6 \in \{q51, q52, q53, q54, q55, q56, q57, q58, q59, q60\} \quad (7)$$

Each player starts with questions in QL_1. Questions arise randomly. Maximum time for each question set to 120 seconds. The next question will be displayed when the player has answered the question, even though the time spent was not until 120 seconds. Start with 3 for life, each correct answer will increase player's life until maximum 5, otherwise decrease player's life. The level of difficulty will increase until game over.

B. Data Collection

For this study, data from 135 normal students and 25 children with special needs, was used. Data were collected randomly among children aged 10 to 12 years in grade 4, 5, and 6 from 2 regular elementary schools and among children with special needs aged 9 to 22 from school for children with special needs. All students were asked to play a math game that had been prepared in the personal computer. Before playing the game, students were given the direction how to play the game. Paper and pen were provided near the computer if students need to do a calculation.

When start playing the game, students fill personal data consisting of name, gender, age, school, and grade. Students play the game until game over. Question id, a level of the game, students' answers, a true or false status of the answer, time spent to answer, and the last stage, are recorded in a game log. Students also completed a written test containing 60 questions, same as the question in the game. Time to complete the written test is 2 hours. Ground truth is determined based on the results of written test and student's mark from the teachers.

Begin
For each S_j
$Check\ if\ (MarkL_1 \geq 7)\ and$
$(MarkL_2 \geq 7)\ and\ (MarkL_3 \geq 7)\ and$
$(MarkL_4 \geq 7) and\ (MarkL_5 \geq 7) and$
$(MarkL_6 \geq 7)\ then\ assign\ L_j = 6$
$Elseif\ (MarkL_1 \geq 7)\ and$
$(MarkL_2 \geq 7)\ and\ (MarkL_3 \geq 7)\ and$
$(MarkL_4 \geq 7) and\ (MarkL_5 \geq 7)$
$then\ assign\ L_j = 5$
$Elseif\ (MarkL_1 \geq 7)\ and$
$(MarkL_2 \geq 7)\ and\ (MarkL_3 \geq 7)\ and$
$(MarkL_4 \geq 7)\ then\ assign\ L_j = 4$
$Elseif\ (MarkL_1 \geq 7)\ and$
$(MarkL_2 \geq 7)\ and\ (MarkL_3 \geq 7)$
$then\ assign\ L_j = 3$
$Elseif\ (MarkL_1 \geq 7)\ and$
$(MarkL_2 \geq 7)\ then\ assign\ L_i = 2$
$Elseif\ (MarkL_1 \geq 7)\ \ then\ assign\ L_j = 1$
$Else\ \ assign\ L_j = 0$
End

Fig. 2. Procedure to define the students' performance level

C. Data Preprocessing

There are 3 data obtained from the data collection i.e. student's mark on the mathematic subject, written test result, and gameplay data.

C.1. Identification of Student Performance Level

In Indonesia, the student is considered mastered the subject at a certain level if the test result at this level meets the minimum standard of a pass. Based on discussion with some experts, a minimum value to pass level in mathematics is 7. This value is used to determine whether the student passes or fail at a certain level. The student is pass in level n if he passes level n and passes the level before n. Written test results are analyzed to get the student performance level of math using 3 steps:

1) Convert status of student answers to the nominal value as (8). a_i is a status of student answer on question i. For each i, set $a_i \leftarrow 0$, if the status of student answer on question i is false and set $a_i \leftarrow 1$ if the status of student answer on question i is true.

$$a_i \leftarrow \{0,1\} \tag{8}$$

2) Add the total value of student answers in each level and put to feature $MarkL_i$ as (9). $MarkL_i$ represents the total value in level i. This feature is extracted based on category [11]. lb is lower bound in each level, and ub is upper bound in each level. lb for level 1, level 2, level 3, level 4, level 5 and level 6 are 1, 11, 21, 32, 41, and 51 respectively. ub for level 1, level 2, level 3, level 4, level 5 and level 6 are 10, 20, 30, 40, 50, and 60 respectively.

$$MarkL_i \leftarrow \sum_{a=lb}^{ub} a_i \tag{9}$$

3) Define the level using the procedure in Fig 2. S_j represents student j.

4) Use student's mark on the mathematical subject and confirm all the student performance levels from step 3 to experts. Use expert judgment as ground truth.

C.2. Gameplay Data Processing

In gameplay data, there are 305 attributes, consisted of 5 attributes of personal data; i.e, name, gender, grade, age, and school and 300 attributes of a game log; consisted of 60 question ids, 60 student answers, 60 question statuses, 60 question levels and 60 time-spent.

Gameplay data is processed to predict math skill level of students. There are 160 student records data with 305 attributes for each record. Before doing classification and prediction, features were selected based on attribute's correlation with class prediction.

Status of student answers in gameplay data was converted to the nominal value and stored to attribute ag_i. ag_i is a status of student answer on question i in the game. For each i, set $ag_i \leftarrow 0$ if the status of the student on question i in the game is false and set $ag_i \leftarrow 1$ if status of student answer on question i in the game is true. Add the total value of student answers in each level as (9) and store it to $MarkGL_1, MarkGL_2, MarkGL_3, MarkGL_4, MarkGL_5$ and $MarkGL_6$ for total value of level 1, 2, 3, 4, 5, and 6 respectively.

Using procedure in Fig 2, performance student levels based on the game were defined and stored in LG_j, j represent the student j. All the predictor variables which were derived from gameplay data are given in Table 1 for reference.

D. Classification and Prediction

Classification process of gameplay data is given in the following.

- Classify gameplay data by applying two test option namely: cross-validation and percentage split. Classification are done in several numbers of folds: 10, 15, 20, 25, and 30 and percentages of split: 70%, 75%, 80%, and 90%.
- Classify gameplay data using 6 data mining methods, i.e. Naive Bayes, MLP, SMO, Decision Table, JRIP, and J48.
- Conduct classification using 9 predictors (all $MarkGLs$, grade, gender, and age), 8 predictors (all $MarkGLs$, grade, gender; all $MarkGLs$, grade, age; all $MarkGLs$, age, gender), 7 predictors (all $MarkGLs$ and grade; all $MarkGLs$ and gender; all $MarkGLs$ and age) and 6 predictors (all $MarkGLs$) for normal students data and children with special needs.
- Analyze classification result and choose a number of fold or percentage of split that gives a maximum percentage of correctly classified instances for the six methods.

TABLE I. STUDENT VARIABLES FOR PREDICTOR FEATURES

Var Name	Description	Domain
Grade	Student's grade	{1,2,3,4,5,6,7,8,9,11,12}
Gender	Student's gender	{1,2}, 1 for male, 2 for female
Age	Student's age	{9,10,11,12,13,14, 15,16,17,18,19,20,21,22}
$MarkGL_1$	Total value of student's answer in game level 1	{0,1,2,3,4,5,6,7,8,9,10}
$MarkGL_2$	Total value of student's answer in game level 2	{0,1,2,3,4,5,6,7,8,9,10}
$MarkGL_3$	Total value of student's answer in game level 3	{0,1,2,3,4,5,6,7,8,9,10}
$MarkGL_4$	Total value of student's answer in game level 4	{0,1,2,3,4,5,6,7,8,9,10}
$MarkGL_5$	Total value of student's answer in game level 5	{0,1,2,3,4,5,6,7,8,9,10}
$MarkGL_6$	Total value of student's answer in game level 6	{0,1,2,3,4,5,6,7,8,9,10}
L	Ground truth for student performance level, derived from written test result and expert judgement	{0,1,2,3,4,5,6}

- Determine the optimum percentage of correctly classified instances for the fold or percentage of split.

IV. RESULT AND DISCUSSION

This section discusses the performance of six different classification methods. Each method was performed for similar gameplay data from 135 normal student data and 25 children with special needs data.

Based on the experiments, we find that the important attributes to predict student performance levels for a normal student are 8 attributes, i.e. age, gender, $MarkLG_1$, $MarkLG_2$, $MarkLG_3$, $MarkLG_4$, $MarkLG_5$, and $MarkLG_6$. Meanwhile, the performance level of children with special needs not influenced by age and grade. Prediction on normal student data also performed well on 6 attributes, i.e. $MarkLG_1$, $MarkLG_2$, $MarkLG_3$, $MarkLG_4$, $MarkLG_5$, and $MarkLG_6$.

Six methods also were implemented on 160 data, normal student data, and children with special needs data, using 6 predictors. Table II shows the comparison of the result at the optimum cross-validation or percentage split test option. SMO gives the best average accuracy with 57.49% accuracy. SMO models are followed by MLP, JRIP, Decision Table, J48, and Naive Bayes with an average accuracy 56.63%, 55.22%, 52.93%, 50.79%, 50.34% respectively. However, JRIP with 10-fold cross-validation produced the best prediction result with accuracy 64.12%.

V. CONCLUSION

In this study, gameplay data was used for predicting math skill level of normal student and children with special needs. Based on experiments, we found that the age, gender, grade, and mark of each level became important factors in determining the level of math skill for the normal student. However, age, gender, and grade don't have a correlation with math level of children with special needs. Six classification methods, Naive Bayes, MLP, SMO, Decision Table, JRip, and J48, were performed to predict math skill performance level of normal students and children with special needs. Based on the classification result, it can be concluded that the best accuracy obtained using JRIP algorithm using 10-fold cross-validation with the highest percentage of accuracy of 64.12.

This study is part of research in developing an intelligent game for assessment of children with special needs. For future, the best algorithm in this study, JRIP, will be used in our game to predict player's level in math skill.

ACKNOWLEDGMENT

The research for this paper was financially supported by Indonesia Endowment Fund for Education (LPDP), Ministry of Finance, Indonesia.

TABLE II. PERFORMANCE OF CLASSIFICATION ALGORITHM

		NB	MLP	SMO	DT	JRIP	J48
Fold	10	52.6	60.3	63.3	61.83	**64.12**	59.54
	15	52.6	58.7	63.3	59.5	60.3	54.96
	20	53.43	59.5	63.3	58.01	59.5	54.9
	25	53.43	61.83	63.3	58.01	58.7	51.14
	30	52.67	56.48	63.3	58.01	61.83	53.43
Percentage split	75%	57.57	60.6	60.6	57.57	60.6	57.57
	80%	50	53,84	57.69	53.8	57.6	53.84
	85%	45	55	50	45	50	45
	90%	38.4	46.15	38.4	30.76	30	30.76
Average Accuracy		**50.34**	**56.63**	**57.49**	**52.93**	**55.22**	**50.79**

REFERENCES

[1] A. M. Shahiri, W. Husain, and N. A. Rashid, "A Review on Predicting Student's Performance Using Data Mining Techniques," *Procedia Computer Science*, vol. 72, pp. 414–422, 2015.

[2] P. Kaur, M. Singh, and G. S. Josan, "Classification and Prediction Based Data Mining Algorithms to Predict Slow Learners in Education Sector," *Procedia Computer Science*, vol. 57, pp. 500–508, 2015.

[3] H. Harwati, A. Permata Alfiani, and F. Ayu Wulandari, "Mapping Student's Performance Based on Data Mining Approach (A Case Study)," *Agriculture and Agricultural Science Procedia*, vol. 3, pp. 173–177, 2015.

[4] V. Ramesh, P. Parkavi, and K. Ramar, "Predicting Student Performance : A Statistical and Data Mining Approach," *International Journal of Computer Application*, vol. 63, no. 8, pp. 35–39, 2013.

[5] B. Şen, E. Uçar, and D. Delen, "Predicting and analyzing secondary education placement-test scores: A data mining approach," *Expert Systems with Applications*, vol. 39, no. 10, pp. 9468–9476, 2012.

[6] M. Qian and K. R. Clark, "Game-based Learning and 21st century skills: A review of recent research," *Computers in Human Behavior*, vol. 63, pp. 50–58, 2016.

[7] Kementerian Pendidikan & Kebudayaan Indonesia, "Executive Summary National Examination 2014 in Indonesia," 2014.

[8] I. N. Sukajaya, I. K. E. Purnama, and M. H. Purnomo, "Intelligent Classification of Learner's Cognitive using Bayes Net, Naïve Bayes, and J48 Utilizing Bloom's Taxonomy-based Serious Game," *International Journal of Emerging Technologies in Learning*, vol. 10, no. 2, pp. 46–52, 2015.

[9] E. Núñez Castellar, J. Van Looy, A. Szmalec, and L. De Marez, "Improving arithmetic skills through gameplay: Assessment of the effectiveness of an educational game in terms of cognitive and affective learning outcomes," *Information Sciences*, vol. 264, pp. 19–31, 2014.

[10] L. Studios, "Monkey tales, in: Die keure & Larian Studios Gent, Belgium," 2011.

[11] Y. Yamasari, S. M. S. Nugroho, I. N. Sukajaya, and M. H. Purnomo, "Features Extraction to Improve Performance of Clustering Process on Student Achievement," in *Proceeding of The 20th International Computer Science and Engineering Conference*, 2016.