EM 11 (1) (2022) 22-26

Educational Management

http://journal.unnes.ac.id/sju/index.php/eduman

The Use of The Learning Analytics Method in Moodle LMS Data to Predict The Final Score of Students in The Vocational Faculty

Andreas Emmanuel Hadisoebroto ✉, **Philipus Suryo Subandoro**

Widya Mandala Catholic University Surabaya, Indonesia

Article Info

Article History:
Recived 01th December 2021
Accepted 02th February 2022
Published 30th April 2022

Keywords:
Moodle, Learning Analytics, liveliness, python, academic achievement

Abstract

The online learning system used by most higher education institutions is moodle. Researchers will also use Moodle as a source of student activity data. Student activity data, including the frequency of accessing the LMS, the number of assignments completed, and the amount of material accessed, can be retrieved and analyzed through data stored in Moodle. Researchers took data in the first and even semesters using the Learning Analytics method combined with statistical analysis and data analytics. The results showed that the submission variable that shows the frequency of completion of students completing assignments or quizzes has a positive influence and a significant correlation to the final score in 1 school year. The duration and action variables both showed insignificant and even negative impacts on students' acquisition of final grades. In addition, the duration and action variables have a solid insignificant correlation to the acquisition of the final value.

✉Correspondence Address:
Vocational Faculty of Widya Mandala Catholic University Surabaya,Agustinus Building, 6th Floor, Dinoyo Street 42-44 Surabaya
E-mail: andreas@ukwms.ac.id

p-ISSN 2252-7001
e-ISSN 2502-454X

INTRODUCTION

Learning Management System, a more popular LMS in higher education, is an online portal connecting lecturers with students. The LMS allows class materials and activities to be shared easily. The portal also enables lecturers and students to interact outside physical classes, discussing using forums that, if done all in physical classes, would consume a lot of time. This online portal should be where students can search and get information about their courses appropriately and reliably (Adzharuddin 2013).

Using this LMS will record all student activities electronically, starting from attendance, assignment work, forums, and discussions. Because this online learning mainly uses the Learning Management System (LMS), variables (Anon n.d.-c) generated by the LMS can be measured. Those variables include how many times the student logs into the LMS, how long the student spends in the LMS, how many times the student accesses a particular module, and the number of variables.

Research using the survey and interview method for students who use google classroom shows that students feel that this technology accelerates the learning process (Subandoro and Sulindra 2019). Learning Analytics (LA) is a method to extract large amounts of learning data available on online learning systems to assist educational institutions in making academic decisions in learning assessment (Puspitasari 2016). This program has been widely used to analyze data on the learning process of students who are members of online classes. The analytical methods used in learning analytics are very diverse according to the data analysis needs. The Data commonly used in Learning Analytics is primarily the characteristics of students, such as their personality and demographics, but much of the recent research has focused on LMS data (Tempelaar, Rienties, and Giesbers 2015).

Learning Analytics can play a significant role in learning by providing information about student activities and their interaction with assignments and online materials so lecturers can reflect on the learning process. Lecturers can use this information as feedback to students, or it can also be used as evidence of changes/improvements to the learning material and/or design. LA can be used as an intervention for all students in a course or to identify students who are at risk of failure (Corrin L, Kennedy G, de Barba P, Lockyer L, Gasevic D, Williams D n.d.).

With increasing demands on the teaching and learning process using technology, the Learning Analytics method can be applied to improve teaching methods by lecturers (Shibani, Knight, and Buckingham Shum 2020). Researchers want to apply LA, which includes methods of measuring, collecting, and analyzing learner behavior (Conijn et al. 2017) on LMS in the faculty Vocational.

The research will also aim to prove a positive relationship and a strong correlation between student activity and academic achievement in online learning. As mentioned above, the variables used in each study are different, with varying degrees of prediction accuracy. This difference can be due to other access rights or the use of various tools in the LMS (Mozahem 2020) so that each institution has a specificity in taking the correct variables. Recent research takes login activity as an indicator that the student is active. From this login activity, it can be concluded that the higher the student's score, the more actively he will log in to the LMS (Mozahem 2020)

METHOD

Learning Analytics allows researchers to use the Data Mining method to collect data using tools to obtain data on the web-based student learning process. Before data processing through Data Mining, several stages of activities are carried out, namely Data Cleansing, Data Manipulation, and Data Wrangling. At the Data Wrangling stage, the Data that is ready to be processed will be combined with statistical analysis methods to get an overview of the conditions of each variable used. In addition to descriptive statistical analysis, regression and correlation analysis were also carried out to determine the relationship and degree of correlation of each existing variable using Pearson's chorus test.

RESULTS AND DISCUSSIONS

The current research focuses on students' personalities as learning participants who use Moodle LMS as an online learning medium. Data processing is carried out using data that has been obtained through the collection of data generated through student interaction with the moodle application used.

The Data is retrieved using the Moodle Configurable Reports plug-in. This plug-in allows limited access to the moodle database so that data can be directly retrieved from the source. The Moodle table used to record the activity of a Moodle user is logstore_standard_log. In this table, all user activities are recorded starting to log in, access courses, access activities in Moodle, and other activities. Data is taken from each stage of UTS and UAS within 1 (one) semester. This Data can calculate how long the user accesses the moodle, judging from the start of the activity until the end. As for the submission of tasks, the Data is taken from the assign_submission table. From this table, data can be obtained on whether students submit assignments or not. All these data will be processed using data analysis methods with multiple regression models used as follows:

$$Y = b0 + b1X_1 + b2X_2 + b3X_3 + b4X_4$$

Where Y is the student's academic achievement, $b_{1\text{-}4}$ is the regression coefficient. $X_{1\text{-}4}$ is the student activity variable consisting of X_1 is the duration of the student log (named the duration variable); X_2 is the number of student log frequencies (named action variables), and X_3 is the number of tasks completed/submitted (named variable submit). Variable X_4 about the amount of material accessed cannot be taken data from the Moodle LMS because not all students and lecturers determine the complete status (*Mark as Done*) for each material in the Moodle LMS. Meanwhile, the dependent variable, namely the student's academic achievement, will be named STS or SAS. Academic achievement data which is the average score of STS and SAS is taken from data for 1 (one) academic year collected by academic sites from Widya Mandala Catholic University Surabaya for each student.

The discussion will be divided on data processing from student activities in the Faculty of Vocational Studies, namely students of the Diploma-3 Office Administration Study Program and students of the Diploma-3 Accounting Study Program.

After the Data is obtained through data mining, data cleansing of each variable is carried out to ensure that each element of the data can be used. Each variable of Data is ensured to have NRP as the Data's primary key to processing. Activities carried out during *Data Cleansing* include *normalizing data*, *Handling Duplicate Data*, *Import* and *Export Data, and Handling Missing Values.* Furthermore, the *Data Manipulation* stage includes activities: *Checking and Changing Data Types, Renaming, Replacing, Removing Data,* and *Filtering Datasets*. The next stage is *Data Wrangling*, which provides for Merging, Transformation, and Statistical Analysis activities.

The processing of correlation data was carried out using an approach from Pearson that can be written in the form of syntax for the first half of STS, as follows:

```
df=pd.read_csv(path+'NA_all_uts.csv')
df_cor=df[['NA','Durasi','Action','submitted']]
df_cor.corr()
```

Meanwhile, the correlation calculations in the second half of the SAS data are as follows:

```
df=pd.read_csv(path+'NA_all_UAS.csv')
df_cor=df[['NA','Durasi','Action','submitted']]
df_cor.corr()
```

The results of the correlation calculation on the first half of the STS and SAS are as follows:

Table 1. Table of Correlation Results of Pearson calculations in the first half of the STS

	Na	Duration	Action	submitted
Na	1.000000	0.442912	0.136428	0.537934
Duration	0.442912	1.000000	0.573603	0.840459
Action	0.136428	0.573603	1.000000	0.508463
submitted	0.537934	0.840459	0.508463	1.000000

Table 2. Correlation Results Table of Pearson calculations first half SAS

	Na	Duration	Action	submitted
Na	1.000000	0.471852	0.196882	0.592723
Duration	0.471852	1.000000	0.547603	0.833020
Action	0.196882	0.547603	1.000000	0.389149
submitted	0.592723	0.833020	0.389149	1.000000

Based on the results of the correlation calculation in table 1 and table 2, it shows that the correlation amount of the variable X_3 (submit) has a reasonably strong correlation value compared to other variables to the value or academic achievement of the Student (NA). Comparing the correlation value of X_3 to NA both in the first half of STS and in the second half of SAS shows that the commitment of students to complete lectures is better than in the first half to get NA grades. Meanwhile, when compared between X_1 (Duration) and X_2 (Action), the correlation value of X_1 (Duration) is relatively more robust than the final score (NA)

In addition to using correlation analysis, a regression test was also carried out using the following syntax:

```
x= NA_all[['Durasi','Action','submitted']]
y= NA_all ['NA']
regr = linear_model.LinearRegression()
regr.fit(x, y)
regr = linear_model.LinearRegression()
regr.fit(x, y)
print(regr.coef_)
```

Then the results obtained for the half-life of the STS are as follows:

[4.67934663e-06 -8.47781290e-04 3.87408204e-01]

and the coefficients in the second half of the US Sare as follows:

[-4.47611655e-06 -1.47247533e-04 4.21336610e-01]

If such coefficients are included in the regression equation, the equation that can be written in the first half of STS is as follows:

$$Y = 4{,}68e\text{-}06X_1 - 8{,}478e\text{-}04X_2 + 0.387X_3 \quad \dots (1)$$

The regression equation in the second half of the SAS is as follows:

$$Y = -4{,}48e\text{-}06X_1 - 1{,}478e\text{-}04X_2 + 0.421X_3 \quad \dots (2)$$

Based on equations (1) and (2), it can be concluded that the existence of variable X_1 (duration) has a positive relationship with the results of NA scores from each student in the first half of STS and turns into a negative association in the second half of SAS. This variable does not always positively influence good or high NA acquisition even though it has a significant relationship with the NA obtained by students.

The severity of the variable X_2 (action) has a negative relationship in both the first half of STS and the second half of SAS to the results of NA scores from each student and showed an insignificant relationship to NA obtained by students.

The severity of the variable X_3 (submit) has a positive relationship in both the first half of STS and the second half of SAS to the results of NA values from each student. This variable shows a strong connection with the variable obtaining NA value compared to other variables.

The data processing results using the Learning Analytics method show that the activeness of completing tasks in the LMS can be used as a determining variable for predicting the success of studies from students.

CONCLUSION

The results of this study support Mozahem's research (2020). Using the variable of activeness using LMS, duration of using LMS, and activeness of completing tasks in LMS cannot be used as a model for predicting study success using LMS. It depends on the digital literacy of learners who use LMS.

ACKNOWLEDGEMENT

This research was funded by an internal research grant from Widya Mandala Catholic University Surabaya. We would like to convey to LPPM of Widya Mandala Catholic University Surabaya for the financial support provided so that this research can be carried out

REFERENCES

Adzharuddin, Nor. 2013. "Learning Management System (LMS) among University Students: Does It Work?" International Journal of E-Education, e-Business, e-Management and e-Learning (January 2013).

Anon. n.d.-a. "Duplicate Labels." Retrieved (https://pandas.pydata.org/pandas-docs/dev/user_guide/duplicates.html).

Anon. n.d.-b. "Merge, Join, Concatenate and Compare."

Anon. n.d.-c. "The Top 40+ Learning Management Systems (LMS) – July 2022."

Conijn, Rianne, Chris Snijders, Ad Kleingeld, and Uwe Matzat. 2017. "Predicting Student Performance from LMS Data: A Comparison of 17 Blended Courses Using Moodle LMS." IEEE Transactions on Learning Technologies 10(1):17–29.

Corrin L, Kennedy G, de Barba P, Lockyer L, Gasevic D, Williams D, et al. n.d. COMPLETING THE LOOP Returning Meaningful Learning Analy : C Data to Learners Project Details.

Malick Sarr. n.d. "How To Detect and Handle Outliers in Data Mining [10 Methods]."

Mozahem, Najib Ali. 2020. "Using Learning Management System Activity Data to Predict Student Performance in Face-to-Face Courses." International Journal of Mobile and Blended Learning 12(3):20–31.

Puspitasari, Kristanti Ambar. 2016. "Utilization of Learning Analytics in Online Learning." Strengthening Teacher Professionalism in Responding to the Challenges of the 21st Century 7(November 2016):205–16.

Shibani, Antonette, Simon Knight, and Simon Buckingham Shum. 2020. "Educator Perspectives on Learning Analytics in Classroom Practice." Internet and Higher Education 46:0–27.

Subandoro, Philipus Suryo, and Eric Sulindra. 2019. "Optimizing Collaborative Learning: Using Google Classroom in Business English Correspondence Class." VOCATIO: Scientific Journal of Administrative And Secretarial Sciences 2(1):46–66.

Tempelaar, Dirk T., Bart Rienties, and Bas Giesbers. 2015. "In Search for the Most Informative Data for Feedback Generation: Learning Analytics in a Data-Rich Context." Computers in Human Behavior 47:157–67.