

A Systematic Literature Review: The Classification of Learning Engagement based on Log File in E-learning (Affective Aspect)

Indra Maulana¹ and Mochamad Bruri Triyono²

^{1,2}Technology and Vocational Education, Yogyakarta State University, Indonesia

Abstract— Most of e-learning platform store the data about act of learning and teaching engagement in a log file. Log file gives us the specific information related to learners’ behavior. A log file is collection or list that comes from user action. The purpose of the research is 1) to improve learning process and students’ learning outcome. Next, 2) to investigate the trend of teaching material that students interest it. The research implies systematic literature review to explore students’ participation based on log file in e-learning. The main contribution of this research is systematic literature review. Based on the previous studies, it had found that the used of e-learning, log file, and learning engagement of affective aspect in many levels of education.

Keywords— e-learning, affective, log file, engagment student.

I. INTRODUCTION

Since Covid-19 is declared as Public Health Emergency of International Concern/PHEIC in all the world by WHO in January 30, 2020 (Susanna, 2020), Indonesia Government announce that the school is close through the issued of Ministry of Education and Culture number ... 2020 (Dewi, 2020). As the consequence, face to face learning system is stopped then learning and teaching activity move from school into home (Handarini, 2020). The policy of school from home has a big impact for the learning effectiveness of Senior High School. The data showed that 1,2 milyar of students must comple to learn from their home because school and campus is close (Habiba et al., 2020). While, UNICEF reported that it’s more than 60 million students and college students in Indonesia which they don’t go to school because of pandemic (Sikirit, 2020). This condition influences the purpose of learning such as online learning facility, teacher competency, the used of media learning, the awareness and resistant of students in online learning (Yusuf, 2017).

Nowadays, e-learning platform appears as the suitable platform for distance learning. This platform connected learning and teaching engagement with technology. E-learning system is not only to provide material and source of learning. But, it helps student to learn by themselves (C.A. Lee et al., 2021). E-learning platform can be used in everywhere and anywhere through internet. Unfortunately, the learning completion is low and the large number of students make the difficulty for teacher (Chiu et al., 2018). E-learning system is one of important characteristics that students can enroll the course any time. Then, students must to finish all of the learning activity at the same time (Y. Lee, 2019). The way of e-learning platform is different than traditional

learning like face-to-face course e that students are expected to finish the progressive lesson in the same time. E-learning system has the openness and flexibility. Other characteristics of e-learning platform are learning environment computer-based. It catches students’ activity without bother them. A log file can record and construct the information such as how did students use the variety of learning source and open the new way to understand students in computer-based environment. In quantitative, a log file can learn students’ behavior and the used of learning content in computer-based (Y. Lee, 2019). Next, the used of correlation analysis in log file to know the impact of students learning activity in e-learning system (Zhang et al., 2020).

Bloom succeeds to introduce the concept of thinking ability that it is called as taxonomy Bloom. This concept divided the aim of education and learning into three domains are cognitive, affective, and psychomotorics (Anderson et al., 2001). Affective domain includes 5 indicators are receiving/attending, responding, valung, organization, and characterization by a value or value complex).

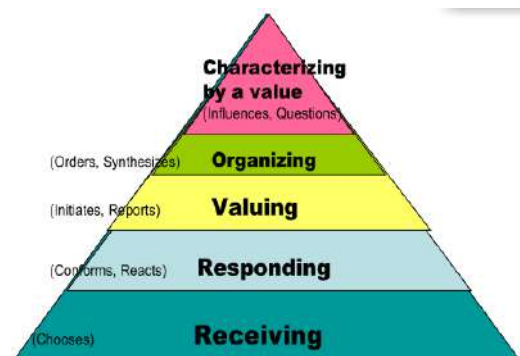


Figure 1.1: Bloom’s Hierarchy of Learning (Allen & Friedman, 2010)

The aim of affective learning includes attitude and students' behaviour that it always be put in syllabus as academic program. The students don't only get the knowledge and skill. But, they need to develop positive attitude and behaviour. However, the previous research of the learning of attitude, appreciation, value, and emotion is less doing by researcher and educator. Some researchers evaluate online learning that it focuses on the aim of cognitive learning than affective learning.

This research observes the factor of students' participation based on log file in the side of e-learning. Researcher focuses on affective aspect in responding. The data of research is collected from Learning Management System (LMS). The data is not only the log file but it is collected from the score total of semester, students' score, assignment score, unfinished assignment, and the total of delaying assignment. A log file shows what kinds of learning content that students click. Furthermore, the variables are courseID, userID, IDcontent click, logTime, exitTime, and IP address. Knowledge discovery process involves data mining that it extracts tendency of data pattern. Then, become the information that it is easy to understand accurately (Han et al., 2012).

The researcher used K-Means algorithm to make cluster of learning engagement. K-Means is cluster algorithm that it arranges based on centroid. The similarity form used distance function. So, it will maximize the similarity of data between short distance and centroid (Han et al., 2012). This research made 3 clusters of students based on log file and the purpose is to see students' participation in e-learning (affective aspect). It focuses on responding indicator. The research questions are 1) Did the e-learning reach students' affective aspect based on log file?, 2) What kinds of learning engagement indicator based on log file?, 3) What kinds of method that it is used to classify learning engagement based on log file?. These questions are conducted because there is no live connection between students and teachers. So, affective aspect doesn't touch yet.

II. RESEARCH OBJECTIVE

This study explores the log file in exploring the active factor of learning using e-learning which focuses on affective aspects. To find out that file logs can contribute to seeing student activity in e-learning, especially in the affective aspect. Appropriate and in-depth research on patterns and how to classify indicators of affective aspects. This study also aims to improve student learning and learning outcomes, and also investigates trends in the teaching materials preferred by students. This study applies for a systematic literature review in

exploring learning activities based on file logs in the e-learning system. The main contribution of this paper is the findings from a systematic literature review based on data found in previous studies, namely the use of e-learning, file logs and also learning in affective aspects at various levels of education.

The SLR results integrate evidence into patterns that can be used to understand state-of-the-art online learning research when applied to the context of various levels of Education. This can provide information to educators to see affective aspects during learning. So that it can find out the activity of students during the learning process, and educators can find out the learning content that is liked by students so that students play an active and enthusiastic role in learning. In addition, conflicting findings from the topic to be analyzed are presented and gaps in the research are highlighted. Section III describes the methods used in SLRs. Section III reports on the SLR results based on the synthesis of evidence. Section IV presents a discussion of the main findings, implications, threats to the validity of this review, and future work.

III. REVIEW METHOD

Systematic Literature Review is defined as the process of identifying, assessing, and interpreting all available research evidence with the aim of providing answers to research questions (Kitchenham et al., 2009). A literature review is an important feature of academic research. Fundamentally, advances in knowledge must build on pre-existing research. To push the boundaries of knowledge, we must know where the limits are (Xiao & Watson, 2019). It is a tool that aims to produce a scientific summary of the evidence in a particular area, as opposed to a "traditional" narrative review. By reviewing the relevant literature, it is possible to identify gaps or gaps to explore that lead to recent research. We followed the procedure of Kitchenham et al (2009).

3.1 The Procedures

3.1.1 Research Questions

The research questions discussed in this study are:

- RQ 1. How to see student learning activities in the e-learning system?
- RQ 2. How to classify student learning activities based on log files?
- RQ 3. How to see students' learning activities on the affective aspect in the log file?

To answer RQ1, we identified studies published in 2018 - 2021, what journals/conferences published them and whether they refer to active learning in e-learning (Chiu et al., 2018,) (Zhang et al., 2020) (Chiu et al., 2020)

(Chiu et al., 2020). al., 2018,) (Zhang et al., 2020) (Chiu et al., al., 2018,) (Zhang et al., 2020) (Chiu et al., 2018). al., 2018)(Aissaoui et al., 2018)(Ray & Saeed, 2018) (Spitzer et al., 2021)(Kuo et al., 2021). With respect to RQ2, we consider the scope of the research (i.e. how to classify learning activities based on log files and using what classification method (Chaudhury & Tripathy, 2018) (Riza et al., 2021) (Sya'iyah et al., 2019) (Maniriho & Effendi, 2018)(Namoun & Alshanqiti, 2021)(Riza et al., 2021)(Kadoic & Oreski, 2018) With respect to RQ3, we considered the individual researcher, the organization to which the researcher is affiliated, and the country to which the organization is located (Phan, 2019) (Wu et al., 2019)(Kaur et al., 2021) (Wei et al., 2021).

The strategy used to find journal articles using PICOC.

Table 1: PICOC

Population	Student,
Intervention	Activities Learning
Comparison	Filelog
Outcame	Affective Aspect
Context	Review dari penelitian tentang keaktifan belajar dalam e-learning dan pembelajaran aspek afektif

3.1.2 Identification of Relevant Literature

The studies used in the search string compilation strategy are as follows (Kitchenham et al., 2009):

- Derive the main terms used in the review question (ie by population, intervention, outcome, and context);
- List the keywords mentioned in the article (main study) that the author knows;
- Find synonyms and alternative words. The study also consulted with the subject librarian for further advice on the proper use of the term;

Highlights that the two main problems in performing SLR searches are the sensitivity and specificity of the search. Sensitivity refers to a search that retrieves a large number of relevant studies. Specificity led to the search for taking the minimum number of irrelevant studies. In the initial search, a small number of articles were retrieved while using the full search string defined above. The keywords “engagement online learning” OR “logfile analysis in e-learning” or “Data Mining Learning Analytics” results in more studies drawn from various online databases. The main search process involved the use of 3 online databases: ScienceDirect, Google Scholar, and Research gate.

After completing the primary search phase, identification of relevant literature is followed by a secondary search phase. During this search phase, all references in the identified paper from the primary source were reviewed. If a paper is found suitable, it is added to the list of eligible existing studies.

3.1.3 Selection of Studies

Search for conference proceedings and specific journal papers that match the research topic since 2017. The journals and conferences selected are in Table 2. The journals were selected because they are known to include literature studies or surveys and have been used as sources for other systematic literature related to software engineering. Each journal and proceeding was reviewed by one of four different researchers (namely Kitchenham, Brereton, and Linkman), and papers dealing with the literature of any type were identified as likely to be relevant. Kitchenham coordinates the allocation of assignments based on the availability of each researcher and their ability to access journals and proceedings at a particular conference. The researcher responsible for searching for a particular journal or conference applies detailed inclusion and exclusion criteria to the relevant paper.

Criteria for research papers 1) study of factors that influence student learning activities. 2) a study that measures student activity 3) a study of the affective domain in learning in e-learning. 4) Study the use of log files. In addition, the following criteria are also applied: 1) Papers presenting supporting evidence claims. 2) a paper describing an Algorithm or Classification method. 3. Papers involving logfile in solving other research topics.

3.1.4 Data Extraction and Study Quality Assessment

To facilitate the extraction process, data formulas were used to gather evidence relating to our question and to measure the quality of the main paper. When designing the paper quality checklist, we reused some of the questions posed in the literature (Fink, 2010). The criteria are based on four quality questions (QA). We assessed studies for quality using the DARE Criteria (Dissemination, 2014). The scores for each study are shown in Table 3. Fields marked with an asterisk in Table 3 were initially marked as unknown and reassigned after communicating with the study authors. Criteria based on four quality questions (QA):

The scoring procedure is Y = 1, P = 0.5, N = 0, or Unknown (ie information is not specified). Kitchenham coordinates the quality evaluation extraction process. Kitchen assessed each paper and allocated 4 papers for

each author of the study to be assessed independently. When there is disagreement, we post issues until we reach an agreement. When a question is scored as unknown, we email the paper's authors asking them to provide relevant information and the question is appropriately scored.

IV. RESULTS AND DISCUSSION

4.1. Search Result

Table 2 shows the search results and we identified 15 specific research articles that match the research topic

that we are going to do based on Scienedirect, scholar, and researchgate. Exploring learning activities in students can take advantage of the wrong log files done by Yaqun Zhang et. al (Zhang et al., 2020). And to analyze student activity in e-learning based on log files, one of which is done by (Kadoic & Oreski, 2018). In the affective domain, learning activity can be seen according to (Phan, 2019). The scores for each research article are shown in Table 2. Each study summarizes its score points.

Table 2: Details of selected articles

ID	Author	Source	Article Title
1	Yu-Chen Chiu ¹ , Hwai-Jung Hsu ¹ , Jungpin Wu ² And Don-Lin Yang (Chiu et al., 2018)	ResearchGate	Predicting Student Performance in MOOCs Using Learning Activity Data
2	Yaqun Zhang (Zhang et al., 2020)	ResearchGate	Using Learning Analytics to Predict Students Performance in Moodle LMS
3	Ouafae El Aissaoui & Yasser El Alami El Madani & Lahcen Oughdir & Youssouf El Allioui (Aissaoui et al., 2018)	ResearchGate	A fuzzy classification approach for learning style prediction based on web mining technique in e-learning environments.
4	Santosh Kumar Ray Khawarizmi, Mohammed Saeed (Ray & Saeed, 2018)	ResearchGate	Applications of Educational Data Mining and Learning Analytics Tools in Handling Big Data in Higher Education
5	Markus Wolfgang Hermann Spitzer, Raphael Gutsfeld, Maria Wirzberger, Korbinian Moeller (Spitzer et al., 2021)	Sciencedirect	Evaluating students' engagement with an online learning environment during and after COVID-19 related school closures: A survival analysis approach
6	Tonny Menglun Kuo, Chin-Chung Tsai, Jyun-Cheng Wang (Kuo et al., 2021)	Sciencedirect	Linking web-based learning self-efficacy and learning engagement in MOOCs: The role of online academic hardness
7	Lala Septem Riza, Rendi Adistya Rosdiyana, Asep Wahyudin, Alejandro	Scholar	The K-Means Algorithm for Generating Sets of Items in Educational Assessment

	Rosales Pérez (Riza et al., 2021)		
8	Khoiriyatus Sya'iyah, Herman Yuliansyah, Ika Arfiani (Sya'iyah et al., 2019)	Scholar	Clustering student data based on K-means algorithms.
9	Pascal Maniriho, Ari Effendi	ResearchGate	Examining the Performance of K-Means Clustering Algorithm
10	Chia-An Lee, Jian-Wei Tzeng, Nen-Fu Huang ¹ and Yu-Sheng Su	ResearchGate	Predicting student performance using data mining and learning analytics techniques: A systematic literature review
11	Nikola Kadoić and Dijana Oreški (Kadoic & Oreski, 2018)	Scolar	Analysis of Student Behavior and Success Based on Logs in Moodle
12	Van Thanh Phan (Phan, 2019)	Scolar	Affective learning objectives in online courses
13	Parneet Kaur *, Harish Kumar, Sakshi Kaushal (Kaur et al., 2021)	Sciendirect	Affective state and learning environment based analysis of students' performance in online assessment
14	Xiaomei Wei, Nadira Saab, Wilfried Admiraal (Wei et al., 2021)	Sciendirect	Assessment of cognitive, behavioral, and affective learning outcomes in massive open online courses: A systematic literature review
15	Moon-Heum Choa Yanghee Kim, DongHo Choia	Sciendirect	The effect of self-regulated learning on college students' perceptions of community of inquiry and affective outcomes in online learning

4.2. Quality evaluation of SLRs

We assessed articles for quality using the DARE Criteria. The scores for each study are shown in Table 2. The results of the quality analysis show that all studies

scored 1 or more on the DA scale and only three studies scored less than 3 (Namoun & Alshantiti, 2021) (Spitzer et al., 2021) (Cho et al., 2017). Two studies scored 4 (Riza et al., 2021) (Kadoic & Oreski, 2018).

Table 3: Quality evaluation of SLRs

Artikel	QA 1	QA 2	QA 3	QA 4	Total Score	Initial rate agreement
1	1	1	0.5	1	3.5	4
2	0.5	0.5	1	1	3	5
3	1	1	1	0	3	4
4	1	0.5	1	1	3.5	3
5	0.5	1	0.5	0.5	2.5	3
6	1	1	0.5	0.5	3	4
7	1	1	1	1	4	5
8	0.5	1	1	0.5	3	5
9	0.5	1	1	1	3.5	3
10	0.5	0	1	0.5	2	3
11	1	1	1	1	4	5
12	1	1	0.5	0.5	3	5
13	1	1	0	1	3	5
14	0.5	1	1	0.5	3	5
15	1	1	0	0.5	2.5	5

The research shows that the log files in the e-learning system can be used to see student learning activities in the affective aspect. Therefore, log files that have been considered only as ordinary data can actually be processed to improve student learning, especially looking at the affective aspects of students. In this study, we can also see that student activity greatly affects learning outcomes.

This study also finds that quantitative research is the researcher's choice. This study suggests that there is a need for further research on affective currents in students' academic achievement. This also proves that the affective domain in learning is very important in learning using the e-learning platform. Interesting content or media for students can also affect student activity

V. CONCLUSION

The findings show that learning activity in e-learning can be seen and can also be an indicator of students' affective aspects in learning. Therefore, there is a need for research that explores the indicators of student activity in the affective aspect. Because learning is not only cognitive and psychomotor, but affective aspects are also very important. Moreover, in distance learning using e-learning, the teacher cannot see directly the activities of students who are safe in learning. Therefore, research on activity based on log files is very important to do.

ACKNOWLEDGMENT

This research is supported by Technology and Vocational Education, Yogyakarta State University, Indonesia.

REFERENCES

- [1] C. A. Lee, J. W. Tzeng, N. F. Huang, and Y. S. Su, "Prediction of Student Performance in Massive Open Online Courses Using Deep Learning System Based on Learning Behaviors," *Educ. Technol. Soc.*, vol. 24, no. 3, pp. 130–146, 2021.
- [2] Y. C. Chiu, H. J. Hsu, J. Wu, and D. L. Yang, "Predicting student performance in MOOCs using learning activity data," *J. Inf. Sci. Eng.*, vol. 34, no. 5, pp. 1223–1235, 2018, doi: 10.6688/JISE.201809_34(5).0007.
- [3] Y. Lee, "Using Self-Organizing Map and Clustering to Investigate Problem-Solving Patterns in the Massive Open Online Course: An Exploratory Study," *J. Educ. Comput. Res.*, vol. 57, no. 2, pp. 471–490, 2019, doi: 10.1177/0735633117753364.
- [4] Y. Zhang, A. Ghandour, and V. Shestak, "Using Learning Analytics to Predict Students Performance in Moodle LMS," *Int. J. Emerg. Technol. Learn.*, vol. 15, no. 20, pp. 102–114, 2020, doi: 10.3991/ijet.v15i20.15915.
- [5] B. Taxonomy, "Bloom's Taxonomy and Digital Tasks," pp. 1–6, 2001.
- [6] L. W. Anderson et al., *A Revision of Bloom's Taxonomy of Educational Objectives*. 2001.

- [7] K. Allen and B. Friedman, "Affective learning: A taxonomy for teaching social work values," *J. Soc. Work Values Ethics*, vol. 7, no. 2, pp. 1–12, 2010.
- [8] B. Kitchenham, O. Pearl Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, "Systematic literature reviews in software engineering - A systematic literature review," *Inf. Softw. Technol.*, vol. 51, no. 1, pp. 7–15, 2009, doi: 10.1016/j.infsof.2008.09.009.
- [9] Y. Xiao and M. Watson, "Guidance on Conducting a Systematic Literature Review," *J. Plan. Educ. Res.*, vol. 39, no. 1, pp. 93–112, 2019, doi: 10.1177/0739456X17723971.
- [10] M. Cocea and S. Weibelzahl, "Can log files analysis estimate learners' level of motivation?" *Lernen, Wissensentdeckung und Adapt. LWA 2006*, no. May 2014, pp. 32–35, 2006.
- [11] O. El Aissaoui, Y. El, A. El, L. Oughdir, and Y. El Alloui, "A fuzzy classification approach for learning style prediction based on web mining technique in e-learning environments," 2018.
- [12] S. Ray and M. Saeed, "Applications of Educational Data Mining and Learning Analytics Tools in Handling Big Data in Higher Education: Trends, Issues, and Challenges Applications of Educational Data Mining and Learning Analytics Tools in Handling Big Data in Higher Education," *Researchgate*, no. July, 2018, doi: 10.1007/978-3-319-76472-6.
- [13] M. W. H. Spitzer, R. Gutsfeld, M. Wirzberger, and K. Moeller, "Evaluating students' engagement with an online learning environment during and after COVID-19 related school closures: A survival analysis approach," *Trends Neurosci. Educ.*, vol. 25, no. July, 2021, doi: 10.1016/j.tine.2021.100168.
- [14] T. M. Kuo, C. C. Tsai, and J. C. Wang, "Linking web-based learning self-efficacy and learning engagement in MOOCs: The role of online academic hardiness," *Internet High. Educ.*, vol. 51, no. April, p. 100819, 2021, doi: 10.1016/j.iheduc.2021.100819.
- [15] P. Chaudhury and H. K. Tripathy, "A Study on impact of smartphone addiction on academic performance," *Int. J. Eng. Technol.*, vol. 7, pp. 50–53, 2018.
- [16] L. S. Riza, R. A. Rosdiyana, A. Wahyudin, and A. R. Pérez, "The k-means algorithm for generating sets of items in educational assessment," *Indones. J. Sci. Technol.*, vol. 6, no. 1, pp. 93–100, 2021, doi: 10.17509/ijost.v6i1.31523.
- [17] K. Sya'iyah, H. Yuliansyah, and I. Arfiani, "Clustering student data based on K-means algorithms," *Int. J. Sci. Technol. Res.*, vol. 8, no. 8, pp. 1014–1018, 2019.
- [18] P. Maniriho and A. Effendi, "Examining the Performance of K-Means Clustering Algorithm," no. March, 2018.
- [19] A. Namoun and A. Alshantqi, "Predicting student performance using data mining and learning analytics techniques: A systematic literature review," *Appl. Sci.*, vol. 11, no. 1, pp. 1–28, 2021, doi: 10.3390/app11010237.
- [20] N. Kadoic and D. Oreski, "Analysis of student behavior and success based on logs in Moodle," 2018 41st Int. Conv. Inf. Commun. Technol. Electron. Microelectron. MIPRO 2018 - Proc., pp. 654–659, 2018, doi: 10.23919/MIPRO.2018.8400123.
- [21] V. T. Phan, "Affective learning objectives in online courses," *Assoc. Comput. Mach.*, pp. 33–36, 2019, doi: 10.1145/3345120.3345189.
- [22] W. H. Wu, H. Y. Kao, S. H. Wu, and C. W. Wei, "Development and evaluation of affective domain using student's feedback in entrepreneurial Massive Open Online Courses," *Front. Psychol.*, vol. 10, no. MAY, 2019, doi: 10.3389/fpsyg.2019.01109.
- [23] P. Kaur, H. Kumar, and S. Kaushal, "Affective state and learning environment based analysis of students' performance in online assessment," *Int. J. Cogn. Comput. Eng.*, vol. 2, no. December 2020, pp. 12–20, 2021, doi: 10.1016/j.ijcce.2020.12.003.
- [24] X. Wei, N. Saab, and W. Admiraal, "Assessment of cognitive, behavioral, and affective learning outcomes in massive open online courses: A systematic literature review," *Comput. Educ.*, vol. 163, p. 104097, 2021, doi: 10.1016/j.compedu.2020.104097.
- [25] A. Fink, *Conducting Research Literature Reviews: From the Internet to Paper*. SAGE Publications, 2010.
- [26] C for R, and Dissemination, "Database of Abstracts of Reviews of Effects (DARE): Quality-assessed Reviews [Internet]," University of York., 2014. <https://www.ncbi.nlm.nih.gov/books/NBK285222/>.
- [27] M. H. Cho, Y. Kim, and D. H. Choi, "The effect of self-regulated learning on college students' perceptions of community of inquiry and affective outcomes in online learning," *Internet High. Educ.*, vol. 34, no. April, pp. 10–17, 2017, doi: 10.1016/j.iheduc.2017.04.001.