

Develop a User Behavior Analysis Tool in ETHOL Learning Management System

¹Dwi Susanto, ²Nuril Ratu Qurani, ³M. Udin Harun Al Rasyid

Politeknik Elektronika Negeri Surabaya (PENS), Surabaya, Indonesia
E-mail: ¹dwi@pens.ac.id, ²nurilratu@it.student.pens.ac.id, ³udinharun@pens.ac.id

Received October 31, 2020; Revised January 19, 2021; Accepted February 9, 2021

Abstract

Students have different learning styles when studying online. Meanwhile, lecturers use the same method for all students who take their online lectures. These different learning styles can affect the level of understanding and the results obtained by students. By knowing student learning styles, lecturers are expected to be able to use the right way in delivering material. In this research, we developed a student behavior analysis feature on self-developed Virtual Learning Environment (VLE) called Enterprise Hybrid Online Learning (ETHOL). Students' data collected includes data on online activities, personal data, and survey data on student learning styles. User behavior analysis was carried out by dividing into three clusters: average scores, time to collect assignments, and student learning styles. The clustering method used is the Hierarchical K-Means. The results obtained are students who have the habit of collecting assignments on time have higher scores than others. In addition, the lecturer is able to see the results of the analysis of the behavior and learning styles of each student. These results can be used as information in delivering lecture material.

Keywords: E-Learning, Personalized Learning, User Behavior Analytics

1. Introduction

There has been a shift in the online learning process, from being instructor-centric to learner-centric. In order to accommodate this shifting, personalization and learning flexibility are enhanced [1]. Personalization allows us to focus on the learning styles that each student likes. By knowing students' learning styles, learning materials can be adjusted to their preferred learning styles. Learning will be more effective for students who have a certain learning style if learning is oriented according to their learning style [2].

There are some of definitions on learning style. One of the definitions explain that learning styles is learning styles are unique behaviors and serve as indicators of how a person learns and how he or she adapts to their environment and how they can provide clues about how someone thinks [3]. Students online learning behaviors can also be used to predict student

performance. These will help developers to evaluate eLearning system effectively and improve system availability [4].

Students' interaction in VLE can be captured in logs. Analyzing these logs will give us information about their achievement in a course, their behaviors and predict their performances [5]. Information on students learning process and interaction can be obtained from their behavior records including learning duration, online assignments, discussion participation, and soon [6]. ETHOL is developed from previous Learning Management System (LMS) [7] and is used in Electronic Engineering Polytechnic Institute of Surabaya (EEPIS). As of Covid-19 pandemic, all its students are pushed to study outside classroom and use ETHOL as a tool for learning. ETHOL record all students' activities during online and we need to provide this information and the analyze result to lecturers. Understanding online students' behaviors will help lecturers to improve their teaching quality and also improve the curriculum.

Based on some of the problems that have been described previously, the authors propose a feature that can help lecturers know the habits and ways of learning from their students. This study proposes an approach using the user behavior analysis method to determine student habits and student learning styles which will then be analyzed so that students' learning style preferences are found. A collection of analyzes for each student will be grouped according to their equations to make it easier for teachers to provide teaching according to the needs and learning styles of students.

2. Related Works

Many researchers have done a lot of research on personalized recommendations such as Hao Li and team [8] who propose methods of recommending E-Learning resources based on learning contexts by constructing students' learning context maps and correlation models of the context of "knowledge resources", combining with personalized recommendation technology, students are treated with learning resources that are appropriate to their learning goals, knowledge abilities and individual preferences. Honghao He and team [9] also proposed a new personalized E-learning service recommendation algorithm proposed to solve the problem of low accuracy, memory and effectiveness. This algorithm builds a matrix of user similarity based on user information data and user behavior data.

In addition, Harish Anantharaman et.al.[10] proposed a comprehensive model based on deep learning by activating the use of Long Short-Term Memory (LSTM) and Random Forest classification, thus providing a highly personalized e-Learning system. Convolutional Neural Networks (CNN) and deep learning algorithms such as Long Short-Term Memory (LSTM) help in identifying learning styles in accordance with Felder and Silverman's learning style modeling (FSLSM).

FSLSM also used by Shaimaa M. Nafea et.al [11] in creating a recommender algorithm to recommends personalized learning object based on students' learning style. FSLSM is used to represent students' learning styles and the learning object profile. To measure the accuracy of the algorithm, the Mean Absolute Error and the Root Mean Square Error are used. Munyaradzi Maravanyika et.al [12] proposed an adaptive recommender system for personalized teaching and learning. The adaptive framework used five dimensions which includes pedagogical strategy, domain model, real-time dynamic adaptation, context modelling, and learner model. Students' learning style based on FSLSM were also analyzed using decision tree and Gradient Boosted tree [13]. Data were collected from questionnaire and records of 65 students in Moodle based LMS which includes: forum, chat, assignments, navigation and course overview.

Learning styles can also be detected based on web usage mining techniques [14]. Instead of asking students to fill out a questionnaire on their learning style, Ouafae El Aissaoui et.al detects students's learning style based on web usage mining. They use clustering algorithms to classify students' log files according to a specific learning style model. The result of clustering are then mapped to the eight categories of FSLM and a learning profile ontology can be modeled. MS Hasibuan et.al [15] used students' prior knowledge as consideration in detecting their learning style. To determine learning style, there are three stages: construct prior knowledge, determine level of knowledge and learning style prediction.

3. Originality

This paper proposes a feature in the virtual learning environment to determine student learning styles. Learning style is obtained based on student behavior data, survey data and assignment-related data. These data are classified, then become the basis for showing student learning styles. These data can also be used to determine student performance.

4. System Design

A general description of the research design in this study can be seen in the following figure 1:

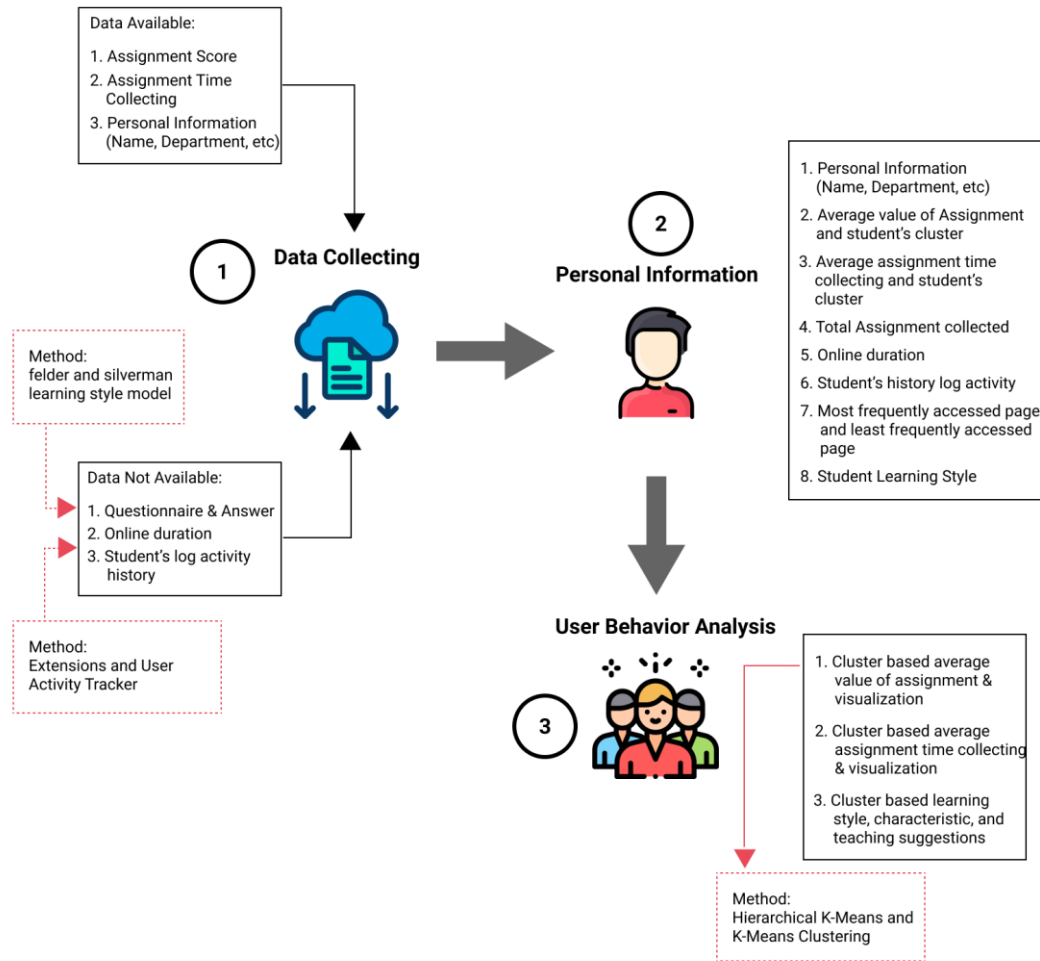


Figure 1. System Design

A system planning and design has been carried out that takes into account the objectives and problems to be solved. This chapter will discuss in detail about the planning and design of the system that the author has made. The general system design is illustrated in Figure 1.

A. Data Collection

Data collection process involved two main process. The first process is to use available data in ETHOL database. The second process is to collect data from students using ETHOL.

Three type of data collected includes:

1. Questionnaire Answers: a survey form in ETHOL filled by students which contain questions related to learning style.
2. Online Time Duration: user time spend during online in ETHOL.
3. Student Log History: this data contains information on students' activities during online in ETHOL.

Two methods are used in data collection:

a. Felder and Silverman Learning Style Model

Students filled out a questionnaire which consist of 44 questions. The questionnaire contains student learning style preferences. Data is processed using Felder and Silverman Learning Styles Model.

Table 1. Learning Style

Active learners are categorized to maintain and understand information better by doing something with learned material such as; discuss, apply or explain it to others.	Reflective students tend to think about concepts quietly first and they like to work alone.
Sensing students prefer to learn facts and learn concrete learning material. To learn from concrete material, students who feel tend to like solving problems with a standard approach and do not like complicated problems.	Intuitive learners are more comfortable with abstract material.
Visual students best remember what they see, such as pictures, diagrams and films.	Verbal learners learn better than written and verbal explanations.
Sequential learners prefer to learn linearly and to find solutions they tend to follow a logical gradual learning path. Sequential learners are more interested in details.	Global learners tend to learn in big leaps and absorb random learning material, they are interested in overviews and find relationships between various fields.

Table I is the process and type of student's learning style using Felder and Silverman Learning Style Model. In the left and right column above explains the differences in learning preferences of each learning style. and what is the most effective way to learn based on each type.

b. User Activity Tracker

Activity Detection Process is an activity to detect events that indicate user activity. Another aspect of the invention is the automatic detection of user activities provided by the use of the User Activity Table. The User Activity Table consists of a list of events that might occur in a system and that show system user activities. The parameters are sufficient to distinguish user activities that indicate events from events that do not indicate usage activities. [16]. User activity tracker includes data login time, logout time, IP address, user full name, activities, information, page names, and scroll.

B. Personal Information

After the data that has been obtained is processed, it will be displayed on personal information. Personal Information is student's personal information which contains:

1. Personal Information includes name, student's id number, and student class.
2. Average Grade Assignments and Clusters. After average assignment grade obtained it will be clustered.
3. Average Time of Assignment and Clusters. students' assignments, besides having grades, also have collection time. The time of collection will be recorded on ETHOL. To find out the tendency of collecting student assignments, an average student collection time must be found. There will be two conditions, on time and late. The collection of assignments on time and late will be added up, and the greatest value will be sought, if more number is on time, then the tendency of the student to collect assignments is on time, and vice versa, if more number of delays, then the tendency of these students is too late in collecting assignment.
4. Total Assignment of Assignments is the total assignments collected by students compared to the total assignments made by lecturers.
5. Duration of Online Time. It is obtained from the logout time of the student which is deducted by the login time or last activity time.
6. Log Activity: A record of the activities carried out by the student on the ETHOL.
7. Most Frequently and Rarely Accessed Pages. Pages that are accessed will be stored in a database, and then will be sorted by frequency of access.
8. Learning Style. It is the type of learning information obtained after students fill out the questionnaire provided. The information includes the process of students receiving an information, processing information to the stage of understanding the information.

C. User Behavior Analysis

User behaviors analysis is personal information collected and grouped according to the equation. The grouping is based on classes and courses that are taught by the lecturer. Distribution based on groups will make it easier for teachers to provide treatment, because it will be troublesome if they have to memorize one by one personal information.

Clustering is divided into three filters:

1. Clusters Based on Average Values and Visualization: After obtaining the average value of each student, the average value of each student will be collected, and will be clustered using the hierarchical k-means method and also k-means. The use of two methods is only intended to compare the accuracy of cluster results. And the method that is really used is hierarchical k-means. There are three clusters in the average value of the assignment, namely the average above which is the group with high grades, the average, and the bottom average which is the

- group with low grades. The grouping will also be visualized using a scatter chart.
2. Clusters Based on Time Collection of Assignments: Will be displayed in the form of pie charts, in the form of a percentage of total students with a tendency to collect assignments on time, and total students with a tendency to collect late assignments. The lecturer will easily infer the tendency of students from the class and provide tailored treatment.
 3. Clusters Based on Learning Styles. Students' learning style will be grouped. There will be two groups in each process, because each process is not related, they will be grouped separately. The perception process will be divided into two group: sensing and intuitive. Input will be divided into visual and verbal, processing will be divided into active and reflective and understanding will be divided into sequential and global. In addition, there are also characteristics and learning suggestions to make it easier for lecturers to provide appropriate treatment and in accordance with the needs of certain group students.

The clustering method uses hierarchical k-means as the main method and k-means as a comparison method.

a. Hierarchical K-Means Clustering

Clustering can be defined as the process of organizing objects in a database into clusters (or groups) so that objects in the same cluster have a high degree of similarity, while objects that are owned by different clusters have a high degree of inequality.

In this research we used Hierarchical K-Means and not ordinary K-Means, because the starting point randomly generated by K-means often makes the grouping results change. Better results from grouping K-means can be achieved after counting more than once. However, it is difficult to determine the calculation boundary, which can give better results. Hierarchical K-Means is applied in order to obtain better results.

b. K-Means Clustering

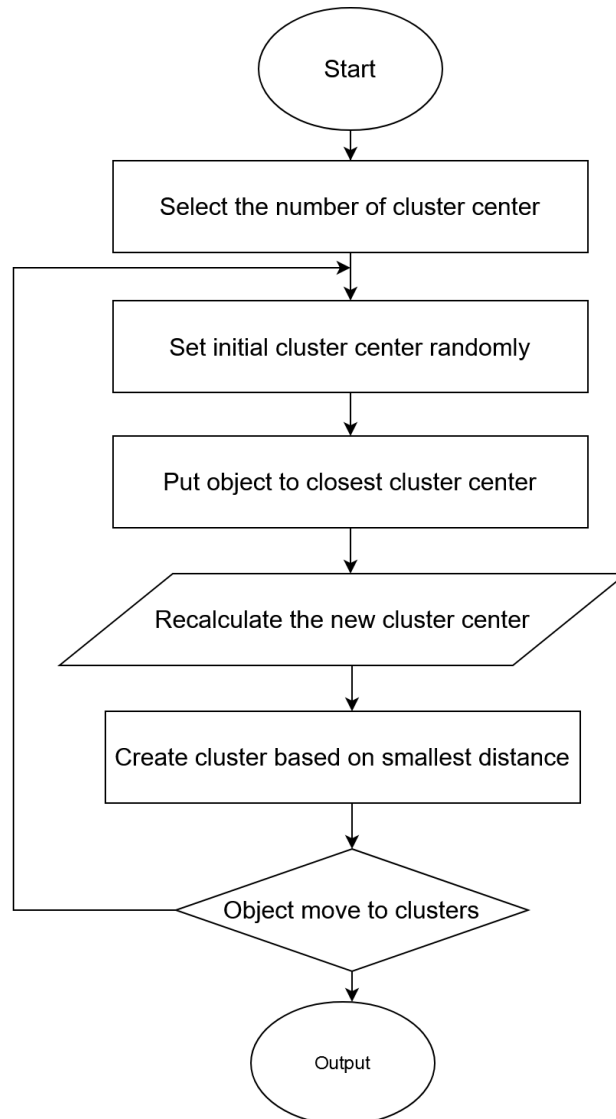
**Figure 2.** K-Means Flow Diagram

Figure 2 explains the K-Means Clustering flowchart. K-means algorithm is one of the most popular grouping algorithms because of its simplicity and fast computation time. K-Means Clustering improves the efficiency of the recommendation system because it can be used off-line to partition data sets in such a way that similar data is grouped in the same cluster and different data are placed in different clusters. Then the recommendation algorithm is applied only to the cluster with the highest similarity. [17]

5. Experiment and analysis

In this research we calculate the accuracy value of k-means and hierarchical k-means and also look for the relationship between student performance and their achievement.

1.1 The accuracy of the K-Means Hierarchy is compared with K-Means

The use of the hierarchical k-means method in this research is because the data in the cluster values do not have training data or are unsupervised data. The clustering process use k-means method.

Trials were conducted twice in different classes, and in one class each test was carried out 10 times each using the hierarchical k-means method and also k-means. The first class has a total of 26 students and the second class has a total of 21 students.

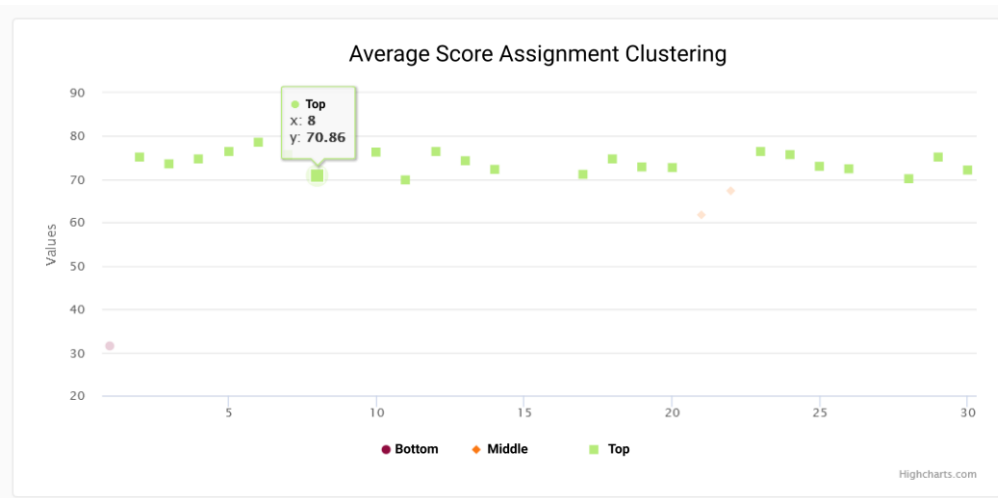


Figure 3. First Class Data Visualization

Figure 3 is the distribution of the average score assignment in the first class. Based on the visualization, there are 23 students in the top cluster, 2 students in the middle cluster and 1 student in the button’s clusters.

Table 2. K-Means and K-Means First-Class Hierarchical Experiments

Experiment	Hierarchical K-Means			K-Means		
	Top	Middle	Bottom	Top	Middle	Bottom
1	13	12	1	13	12	1
2	23	2	1	21	4	1
3	19	6	1	13	12	1
4	13	12	1	13	12	1
5	19	6	1	13	12	1
6	23	2	1	23	2	1
7	23	2	1	16	9	1
8	23	2	1	19	6	1
9	23	2	1	23	2	1
10	19	6	1	23	2	1
Result	5/10 or 0,5			3/10 or 0,3		

Based on the table II, for 10 experiments conducted using the hierarchical k-means method, it has 5 times the same results, whereas when using the k-means method have 3 times the same cluster results.

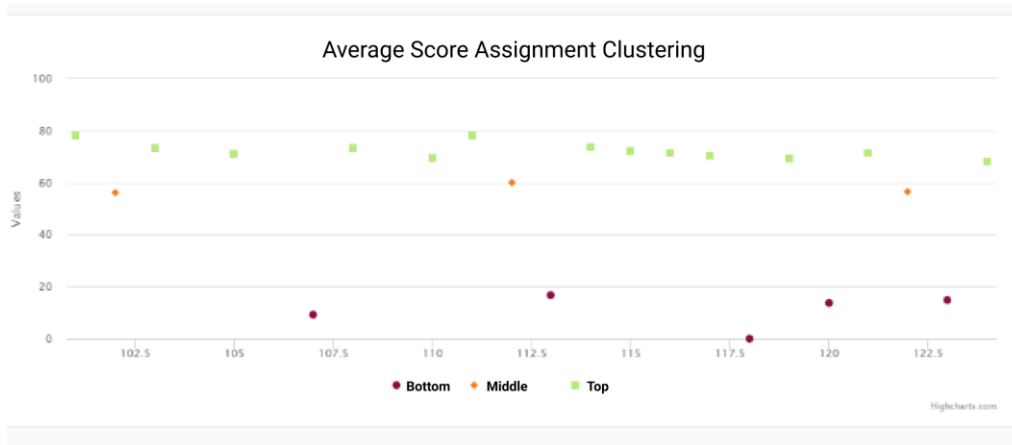


Figure 4. Visualization of Second Class Data

Figure 4 contain the distribution of the average score assignment in the second class. Based on the visualization there are 13 students in the top cluster, 3 students in the middle cluster and 5 students in the button's clusters.

Table 3. Second K-Means and K-Means Hierarchical Experiments

Experiment	Hierarchical K-Means			K-Means		
	Top	Middle	Bottom	Top	Middle	Bottom
1	13	3	5	13	3	5
2	13	3	5	13	3	5
3	13	3	5	13	3	5
4	13	3	5	13	3	5
5	13	3	5	13	3	5
6	13	3	5	12	4	5
7	13	3	5	13	3	5
8	13	3	5	13	3	5
9	13	3	5	13	3	5
10	13	3	5	13	3	5
Result	10/10 or 1,0			9/10 or 0,9		

Table III contain data from the second class for 10 times experiments conducted using the hierarchical k-means. This method has 10 times the same results, whereas when using the k-means method have 9 times the same results.

The accuracy value obtained by hierarchical k-means is higher than k-means. The cause of the accuracy value of hierarchical k-means is higher than k-means is due to the repetition and evaluation of the appearance of

centroid values that are randomly raised. In the hierarchical k-means centroids appear as many entries as possible from the user and then the centroid will be measured using a single linkage and will be evaluated until the centroid amounts to user input.

Table 2 and Table 3 shows the comparison. The distribution of values also affects the performance of the k-means method and also the hierarchical k-means. The farther the distance between data clusters, the higher the accuracy of the k-means and hierarchical k-means methods, whereas conversely the closer the distance between clusters, the lower the accuracy of the k-means and hierarchical k-means methods. The distribution of values will also affect computing time.

Table 4. K-Means and Hierarchical K-Means Time Performing

	Time	Size
K-Means	18,53 s	2,28 KB
Hierarchical K-Means	23,21 s	2,25 KB

However, although the accuracy value of hierarchical k-means is higher than k-means, hierarchical k-means has a longer computational time compared to k-means.

1.2 Relationship between Performance and Student Achievement

Table 5. First Class Student Performance

Name	Average Online Time	Task Collection Time	Total Task Collection	Avg Value
Nur Syaela Majid	63 minutes	On time	8/8	(78) Upper Avg
Achmad Sofyan	117 minutes	On time	8/8	(73,2) Upper Avg
Devvana Arya Pratama	36 minutes	On time	7/8	(56) Medium Avg
Risalatun Nuriyah	60 minutes	On time	8/8	(73,2) Upper Avg
Irfan Hidayah Tullah	36 minutes	On time	8/8	(71,1) Upper Avg
M. Bagas Asrori	48 minutes	On time	8/8	(69,4) Upper Avg
Gray Cloudy Akbar	9 minutes	Late	2/8	(16,9) Lower Avg

Table 6. Second Class Student Performance

Name	Average Online Time	Task Collection Time	Total Task Collection	Avg Value
Akta Raharja Bahrul	39minutes	On time	7/7	(71,14) Medium Avg
Dizca Rosalia Putri	51minutes	On time	7/7	(70,86) Medium Avg
Rizqi Hammaz Zacky	63minutes	Late	5/7	(31,57) Lower Avg
Musayadah Khusnul K	45minutes	On time	7/7	(75,71) Upper Avg
Mahdanu Oktarinda F	60minutes	On time	7/7	(76,29) Upper Avg
Anfa'ul Fitri	99minutes	Late	7/7	(76,43) Upper Ag
Muhammad Yasin	57minutes	On time	7/7	(73,57) Upper Avg

Table 5 and Table 6 shows that students with a group of top grades have a habit of collecting all assignments given by the instructor, and most of them submit assignments on time. Whereas for the average underclass students, they have different habits, which have a habit of rarely collecting assignments, and more often being late when collecting assignments.

Students with middle average scores they have a habit that is sometimes similar to lower average students and sometimes similar to upper average students.

6. Conclusion

This research stores student learning information, including accuracy in the collection of assignments, online duration, student habits on ETHOL and grades from students. In addition, it also adopts Felder and Silverman Learning Style models that are used to find the learning styles of each student. The information stored is grouped based on its similarity to make it easier for teachers to provide useful references to help provide learners in accordance with student learning styles.

How to learn from one student to another is different. Student's performance will affect the results obtained. Special treatment is needed for each student. The difference in learning styles will affect the learning process, and it will affect student achievement. The role of instructors is also needed to provide teaching in accordance with student preferences in order to provide maximum results.

In the next research, it is hoped that it can add records for other activities to be analyzed in order to make it easier to find out the habits of students. Apply the results of learning style questionnaire in teaching and learning activities, and then compare the results before application and afterwards to determine the effectiveness of the learning style.

REFERENCES

- [1] Dongming Xu, Wayne W. Huang , Huaiqing Wangd, Jon Heales a, **Enhancing e-learning effectiveness using an intelligent agent-supported personalized virtual learning environment: An empirical investigation**, *Elsevier, Information & Management*, Vol. 51, Issue 4, pp.430-44, 2014.
- [2] Gunathilaka T. M. A. U, Fernando M. S. D, **Identification of the Learning Behavior of the Students for Education Personalization**, International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud), 2017.
- [3] Ismail El Haddioui, Mohamed Khaldi, **Learning Style and Behavior Analysis A Study on the Learning Management System Manhali**, *International Journal of Computer Applications*, Vol. 54, No. 4, 2012.
- [4] Wei Zhang, Xujun Huang, Shengming Wang, Jiangbo Shu, Hai Liu, Hao Chen, **Student Performance Prediction via Online Learning Behavior Analytics**, International Symposium on Educational Technology, 2017.
- [5] André Luiz de Brandão Damasceno, Dalai dos Santos Ribeiro, Simone Diniz Junqueira Barbosa, **What the Literature and Instructors Say about the Analysis of Student Interaction Logs on Virtual Learning Environments**, *IEEE Frontiers in Education Conference (FIE)*, 2019.
- [6] Baolin Yi, Yi Wang, Dujuan Zhang, Hai Liu, Jiangbo Shu, Zhaoli Zhang, Yuegong Lv, **Learning Analytics-Based Evaluation Mode for Blended Learning and Its Applications**, International Symposium on Educational Technology (ISET), 2017.
- [7] Pernanda, S.T., Winarno I., Susanto, D., Yuwono, W, **Design and Development of E-Learning Mobile Based Interface (Case Study: DOSENJAGA PENS)** , International Electronics Seminar EEPIS, 2014.
- [8] Hao Li, Libin Wang, Xu Du, Mingyan Zhang, **Research on the Strategy of E-Learning Resources Recommendation based on Learning Context**, The Sixth International Conference of Educational Innovation through Technology, 2017.
- [9] Honghao He, Zhengzhou Zhu, and Qun Guo, **A Personalized E-learning Services Recommendation Algorithm Based on User Learning Ability**, *IEEE 19th International Conference on Advanced Learning Technologies (ICALT)*, 2019.
- [10] Harish Anantharaman, Abdullah Mubarak, B.T Shobana, **Modelling an Adaptive e-Learning system using LSTM and Random Forest**

- classification**, IEEE Conference on e-Learning, e-Management and e-Services (IC3e), 2018.
- [11] Shaimaa M. Nafea, François Siewe, Ying He, **A Novel Algorithm for Course Learning Object Recommendation Based on Student Learning Styles**, International Conference on Innovative Trends in Computer Engineering (ITCE), 2019.
- [12] Munyaradzi Maravanyika, Nomusa Dlodlo, Nobert Jere, **An adaptive recommender-system based framework for personalised teaching and learning on e-learning platforms**, IST-Africa Week Conference (IST-Africa), 2017.
- [13] Yunia Ikawati, M. Udin Harun Al Rasyid, Idris Winarno, **Student Behavior Analysis to Detect Learning Styles in Moodle Learning Management System**, International Electronics Symposium (IES), 2020.
- [14] Ouafae El Aissaoui, Yasser El Madani El Alami, Lahcen Oughdir, Youssef El Alloui, **Integrating web usage mining for an automatic learner profile detection: A learning styles-based approach**, International Conference on Intelligent Systems and Computer Vision (ISCV), 2018.
- [15] M S Hasibuan, L E Nugroho, PI Santosa, **Detecting Learning Style Based on Level of Knowledge**, Third International Conference on Informatics and Computing (ICIC), 2018.
- [16] Johnson et al, **Method of And Apparatus for Providing Automatic Detection of User Activity**, International Business Machines Corporation CP Command and Utility Reference, Release 1 IBM No. SC24-5519-00, 1994.
- [17] K. Alhasan, L. Chen, F. Chen, **Mining Learning Styles for Personalised E-Learning**, IEEE Smart World, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovations, 2018