

Ease Evaluation Using the Best Moodle Learning Management System with Data Mining Concepts

Harry Dhika¹, Fitriana Destiawati², Michael Sonny³, Surajiyo⁴

^{1,2,3,4} Indraprasta PGRI University, Jakarta, Indonesia

¹ dhikatr@yahoo.com, ² honeyzone86@gmail.com, ³ michael.sonny@unindra.ac.id,
⁴ surajiyo@unindra.ac.id

Abstract: Education is now integrated with technological developments, especially developments towards the era of industrial revolution 4.0, this is seen based on the many popular Learning Management Systems (LMS) including Moodle. Through this study, searching for the best LMS with the concept of data mining uses the Knowledge Discovery in Database (KDD) method which is applied to students in the Yogyakarta area as many as 246 people. Moodle devices that were made were tested and given to schools with attribute testing, used as many as 16 attributes including the appearance of Moodle which was considered interactive or not, speed of application access, features possessed by Moodle to the user's final conclusion. Evaluation using the Receiver Operating Characteristic (ROC) curve and looking at the value graph with Area Under Cover (AUC). Data is processed with Rapidminer software with 9 KDD steps. From all data, empirical data testing will be conducted so that a model from Moodle LMS is formed. The results of the study show that the accuracy value is above 90%, indicating that the technique of applying data to software is categorized very well. So that the main focus in determining LMS is said to be the best can be validated and ascertained accurately in calculating data retrieval from teachers, students and school equipment as the object of research.

Keywords: learning management system Moodle, mining data, evaluation Moodle knowledge, education, learning

INTRODUCTION

The industrial revolution is now present in our midst, the industrial revolution 4.0 (Fatmawati, 2018). Education is a field that is definitely affected and inevitably enters the industrial revolution. On 29 January 2019, the Ministry of Research and Technology has actively conveyed the industrial revolution in the field of education. Ministry of Research and Technology delivered a program of globalization of education and industrial revolution 4.0 (Ghufron, 2018). Through this matter, the world of education through e-learning is the main answer. Elearning is one part of the industrial revolution 4.0, (Yustanti & Novita, 2019) developing learning makes teachers, students (Kosasi, 2015) and parents easier in the learning process.

The background of this study begins with the need for education for learning in schools (Yunis & Telaumbanua, 2017). At present, with rapid technological developments accompanied by the many gadgets in the school environment and used by every child, it is hoped that it will be used to simply play using gadgets but is used for the purposes of learning and education, the implementation of learning using elearning.

The platform used is the learning management system (LMS) (Raharja, Prasojo, Ariyawan, & Nugroho, 2011), evaluating the use of elearning, especially on Moodle. The study was conducted in Yogyakarta Special Region, Moodle LMS became the main study because of its ease and superiority which is very familiar to teachers or students, because moodle is included in CAL or CAT namely (Computer Assisted Learning + Computer Assisted Teaching), this is a major component in learning and teaching. The design and development of Moodle is driven by

a philosophy of learning. A way of thinking that a person is in the social development theory (social constructionist education). Some scientists have put forward the idea of "soft educational mumbo jumbo" where one simply uses the mouse to obtain education (Raharja et al., 2011). Moodle or Modular Object Dynamic Learning Environment is a simple web application that is used for online learning, learning to use web pages on a computer or being able to use handsets in a class both planned by first preparing a computer device or directly with an existing device.

Entities in moodle can be adapted to needs and not difficult in access settings. Entities made consist of administrators, teachers, students and parents of students. The administrator ensures that there are no problems and prepares all the main systems and configures the moodle, the teacher entity like a teacher, can do the teaching by uploading material documents and making practice questions, the results of the training provide grades or grades that can be used for daily examinations up to semester exam at school. Student entities are objects that do or do the learning process, read material from the teacher and interact with the teacher as befits conventional learning. Students can communicate with the teacher, such as the teacher discussing with other teachers or with the admin. There is also a communication room for students. Students download the material and work on the practice questions provided by the teacher, students can see the material, study it well and be ready to take the scheduled exam prepared by the teacher. The next entity is the parents of students with access to see the value and presence of students who can evaluate learning at home by paying attention to student learning patterns.

The importance of measurements to determine the ease of use of LMS Moodle is due to the rapid development of Moodle LMS compared to other LMS. Classes can be done online or offline with a network at school, this is the main point, moodle is very easy to use with a simple and powerful display, moodle can be obtained for free just download it on the moodle web only. This is a very interesting offer with links to industry 4.0.

METHOD

This research was conducted at several schools in the city of Yogyakarta with a sample of 246 students, teachers, IT staff and some of the parents of students or guardians of students. Processing data using Rapidminer software (RapidMiner, 2014). With the concept of data mining, it is looking at data extraction and finding new data patterns that will be interpreted so as to provide a special pattern that can be used and utilized for the development of Moodle LMS, specifically evaluating and developing for the benefit of the school. The research method uses Knowledge Discovery in Database (KDD) consisting of 9 stages (Maimon, 2010) with 1. Domain Understanding 2. Selection and Addition 3. Preprocessing: Data Cleaning 4. Transformation, stages 5, 6, and 7. Data Mining 8. Final Evaluation and Interpretation 9. Discovered Knowledge (Visualization and Integration). As in figure 1.

After finding and processing data the next activity is to evaluate the data using the Receiver Operating Characteristic (ROC) curve and see the value graph using the Under Curve Area (AUC). Like table 1.

The data mining algorithm used is the C4.5 decision tree with a very popular classification technique that produces rules. Data in numeric and discrete form can be processed by the C4.5 algorithm, from the generated results there are several missing attributes, this can degenerate according to the pruning decision with the value of gain and the entropy value. All rules are very easy to understand and interpret quickly compared to other algorithms. Speed and time given are very short in processing large data.

In calculating the C4.5 algorithm (Sugianto, 2015)(Phadikar, Sil, & Das, 2013) there are stages including:

1. Preparing data to be processed with the available number of 246 records
2. Determine the value of entropy with the formula:

$$Entropy(S) = \sum_{i=1}^n -p_i \cdot \log_2 p_i \dots\dots\dots(1)$$

Information:

S = set of cases

n = number of partitions S

p_i = the proportion of S_i to S

3. Calculate the Gain value as follows:

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \dots\dots\dots(2)$$

Information:

S = set of cases

A = feature

n = the number of partitions of attribute A

$\frac{|S_i|}{|S|}$ = the proportion of S_i to S

|S| = number of cases in S

4. Next do the repetition in step 2 until all processes are completed
5. The process will stop when:
 - a. All records in node N will have the same class
 - b. There are no records that can be counted and partitioned again
 - c. In the branch there is no record

Table 1. Confusion Matrix Model(Milinković & Maksimović, 2013; Zareiforush, Minaei, Alizadeh, Banakar, & Samani, 2016)

Correct Classification	Classified as	
	+	-
+	True Positive	False Negative
-	False Positive	True Negative

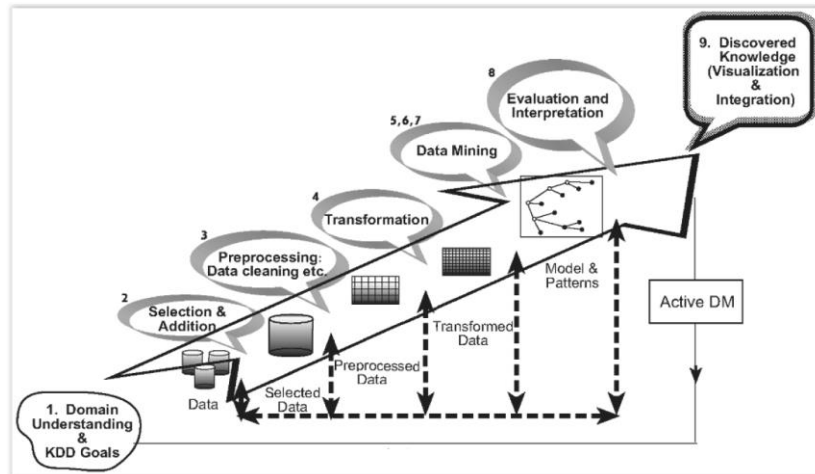


Figure. 1. Knowledge Discovery in Database (Maimon, 2010)

RESULTS AND DISCUSSION

The design carried out with the Rapidminer 5.3(RapidMiner, 2014) application, there are several attributes that are used as follows:

Table 2. Data Attributes and Statistics

Role	Attribute	Type	Statistics	Range
Label	Final Qualification	Binominal	mode = Good (231), least = Bad (14)	Good (231), Bad (14)
Regular Display		Polynomial	mode = Interactive (114), least = Interactive enough (49)	Interactive (114), Less interactive (82), Interactive enough (49)
Regular Speed		Binominal	mode = 0-10 sec (189), least = 10-20 sec (56)	0-10 sec (189), 10-20 sec (56)
Regular Feature		Binominal	mode = Complete (189), least = Incomplete (56)	Complete (189), Incomplete (56)
Regular Mobile Support		Binominal	mode = Support (169), least = Does not support (76)	Support (169), Does not support (76)
Regular	Supports Multiple Formats	Binominal	mode = Many Formats (222), least = Not too much (23)	Many Formats (222), Not too much (23)
Regular	Management Module	Binominal	mode = There is (203), least = Not much (42)	There is (203), Not much (42)
Regular	Support Assignment	Binominal	mode = There is (171), least = There is no (74)	There is (171), There is no (74)
Regular	Internal Communication	Binominal	mode = Support (239), least = Does not support (6)	Support (239), Does not support (6)
Regular	Learning Catalog	Binominal	mode = There is Interactive (185), least = Not interactive (60)	There is interactive (185), Not interactive (60)
Regular	Tracking Learning Activities	Polynomial	mode = There is (143), least = There is no (48)	There is (143), There is no (48), Not known (54)

Role	Attribute	Type	Statistics	Range
Regular Assessment		Binominal	mode = Flexible (204), least = Inflexible (41)	Flexible (204), Inflexible (41)
Regular Reporting System		Binominal	mode = There and easy (201), least = There and difficult (44)	There and easy (201), There and difficult (44)
Regular Security		Binominal	mode = Available (242), least = Not available (3)	Available (242), Not available (3)
Regular Webinar		Binominal	mode = Support (189), least = Not Supported yet (56)	Support (189), Not Supported yet (56)
Regular Infographics		Binominal	mode = Available (245), least = Available (245)	Available (245)

The data sample used including the label is 246 records. The entire data is broken down into two parts training and testing with training composition of 80% and testing by 20%. All training data will be tested and testing data used to test training data that has produced a decision tree model.

The calculation produces the decision tree as follows:

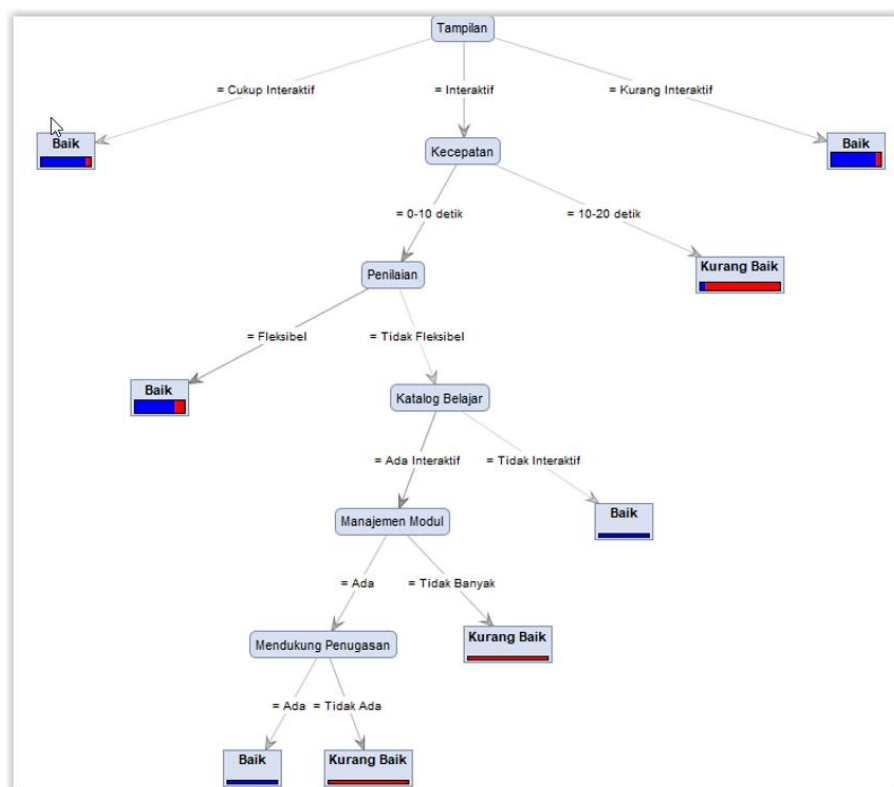


Figure. 2. Result Decision tree

With the definition as follows:

Display = Quite Interactive: Good {Good = 45, Bad = 4}
 Display = Interactive
 | Speed = 0-10 seconds
 | | Rating = Flexible: Good {Good = 54, Bad = 12}
 | | Rating = Not flexible

```

|   |   |   Learning Catalog = There is Interactive
|   |   |   |   Module Management = Yes
|   |   |   |   |   Support Assignment = Yes: Good {Good = 2, Bad = 0}
|   |   |   |   |   Support Assignment = None: Bad {Good = 0, Bad = 3}
|   |   |   |   |   Module Management = Not Much: Bad {Good = 0, Bad = 5}
|   |   |   |   Learning Catalog = Not Interactive: Good {Good = 2, Bad = 0}
|   |   |   |   Speed = 10-20 seconds: Bad {Good = 3, Bad = 33}
|   |   |   |   Display = Not Interactive: Good {Good = 75, Bad = 7}

```

Has conditions:

- IF Display = Simply Interactive THEN Good.
- IF Display = Interactive THEN IF Speed = 0 - 10 seconds THEN IF Flexible Rating THEN Good.
- IF Display = Interactive THEN IF Speed = 0 - 10 seconds THEN IF Assessment Not Flexible THEN IF Learning Catalog = There is Interactive THEN IF Management Module = THEN IF Supports Assignment = There is a Good THEN.
- IF Display = Interactive THEN IF Speed = 0 - 10 seconds THEN IF Assessment Not Flexible THEN IF Learning Catalog = There is Interactive THEN IF Management Module = THEN IF Supports Assignment = No THEN Not Good.
- IF Display = Interactive THEN IF Speed = 0 - 10 seconds THEN IF Rating Not Flexible THEN IF Learning Catalog = There is Interactive THEN IF Management Module = Less Much THEN Less Good.
- IF Display = Interactive THEN IF Speed = 0 - 10 seconds THEN IF Rating Not Flexible THEN IF Learning Catalog = Not Interactive THEN Good.
- IF Display = Interactive THEN IF Speed = 10 - 20 seconds THEN IF Flexible Rating THEN Poor.
- IF Display = Less Interactive THEN Good.

The accuracy measurement results for the evaluation data are as follows:

PerformanceVector:

accuracy: 93.90% +/- 3.29% (micro: 93.88%)

ConfusionMatrix:

True:	Good	Bad
Good:	225	9
Bad:	6	5

precision: 45.45% (positive class: Bad)

ConfusionMatrix:

True:	Good	Bad
Good:	225	9
Bad:	6	5

recall: 40.00% +/- 48.99% (micro: 35.71%) (positive class: Bad)

ConfusionMatrix:

True:	Good	Bad
Good:	225	9
Bad:	6	5

AUC (optimistic): 1.000 +/- 0.000 (micro: 1.000) (positive class: Bad)

AUC: 0.961 +/- 0.030 (micro: 0.961) (positive class: Bad)

AUC (pessimistic): 0.922 +/- 0.061 (micro: 0.922) (positive class: Bad)

The results of the measurement of data classification with an accuracy rate of 93% or 0.93 are included in the following groups which are classified as very good (Gorunescu, 2011; Nancy Jane, Khanna Nehemiah, & Arputharaj, 2016).

- 0.90-1.00 = classification is very good
- 0.80-0.90 = good classification
- 0.70-0.80 = enough classification
- 0.60-0.70 = bad classification
- 0.50-0.60 = wrong classification

The measurement results use the ROC curve as follows:

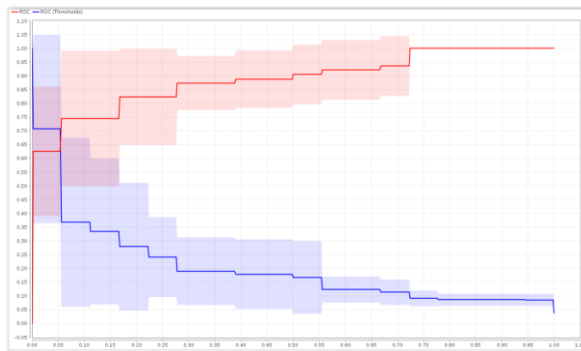


Figure 3. Result Area Under Curve (AUC) Optimistic

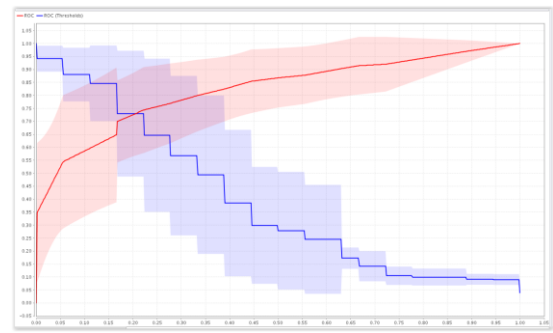


Figure 4. Result Area Under Curve (AUC) Positive Class

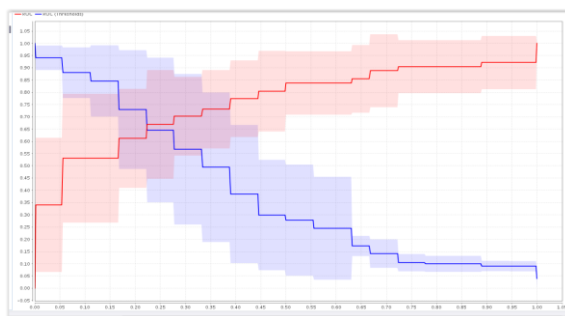


Figure 5. Result Area Under Curve (AUC) Pessimistic Class



Figure 6. Moodle Dashboard

In figure 3. Determining the value of the Optimistic Under Curve Area, with optimistic values reaching 0.9 and on a scale of 1 maximum. Figure 4. shows the AUC Positive class value up to 0.961 giving information that the classification is very good with a constant graph leading and focusing on the top of the graph section, a positive plot case before seeing the Negative example. Figure 5. Shows the Pessimistic Class value in the case of a negative plot before seeing a positive example produced. All prediction sequencing is done to determine the best value of the AUC Optimistic and Pessimistic average.

CONCLUSION

This study discusses the evaluation of the ease of using LMS Moodle in Senior High Schools in Yogyakarta, where moodle is a LMS that is very widely used and powerful, for that it obtained satisfactory results in this study namely the accuracy of data up to 93% with 246 data including labels and 16 attributes that are not related to each other. The method used is 9 steps KDD, evaluation using the ROC curve and measuring data precision and accuracy of the data, with high data precision values and high accuracy 93%, it can be concluded that the classification category is very good.

REFERENCES

- Fatmawati, E. (2018). Disruptif diri pustakawan dalam menghadapi era revolusi industri 4.0. *Iqra'*, 12(01), 1–13. Retrieved from jurnal.uinsu.ac.id/index.php/iqra/article/download/1816/1479%0A
- Ghufron, M. A. (2018). Revolusi industri 4.0: Tantangan, Peluang dan Solusi Bagi Dunia Pendidikan. *Seminar Nasional Dan Diskusi Panel Multidisiplin Hasil Penelitian Dan Pengabdian Kepada Masyarakat*.
- Gorunescu, F. (2011). Data mining: Concepts, models and techniques. *Intelligent Systems Reference Library*, 12. <https://doi.org/10.1007/978-3-642-19721-5>
- Kosasi, S. (2015). Perancangan E-learning untuk Meningkatkan Motivasi Belajar Guru dan Siswa. *Jurnal Informatika*, (0362), 27213. <https://doi.org/10.1007/s10619-011-7079-6>
- Maimon, O. (2010). *Data Mining And Knowledge Discovery Handbook* (Second Edi; Prof. Oded Maimon, Ed.). <https://doi.org/10.1007/978-0-387-09823-4>
- Milinković, S., & Maksimović, M. (2013). Using Decision Tree Classifier for Analyzing Students ' Activities. *Jita*, 3(2), 87–95. <https://doi.org/10.7251/JIT1302087M>
- Nancy Jane, Y., Khanna Nehemiah, H., & Arputharaj, K. (2016). A Q-backpropagated time delay neural network for diagnosing severity of gait disturbances in Parkinson's disease. *Journal of Biomedical Informatics*, 60, 169–176. <https://doi.org/10.1016/j.jbi.2016.01.014>
- Phadikar, S., Sil, J., & Das, A. K. (2013). Rice diseases classification using feature selection and rule generation techniques. *Computers and Electronics in Agriculture*, 90, 76–85. <https://doi.org/10.1016/j.compag.2012.11.001>
- Raharja, S., Prasajo, L. D., Ariyawan, D., & Nugroho, A. (2011). Model Pembelajaran Berbasis Learning Management System Dengan Pengembangan Software Moodle Di Sekolah Menengah Atas Learning Model Based Learning Management System With Software Development Based Moodle At Seior High School. Universitas Negeri Semarang.
- RapidMiner. (2014). RapidMiner Studi o Manual. In G. RapidMiner (Ed.), *RapidMiner* (6th Editio). <https://doi.org/http://docs.rapidminer.com/downloads/RapidMiner-v6-user-manual.pdf>
- Sugianto, C. A. (2015). Penerapan Teknik Data Mining Untuk Menentukan Hasil Seleksi Masuk Sman 1 Cibeber Untuk Siswa Baru Menggunakan Decision Tree. *TEDC*, 9, 39–43. Retrieved from <https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&cad=rja&uact=8&ved=2ahUKEwj11- jyu97jAhVHrY8KHYYJ1B2wQFjABegQIAhAC&url=%2Finterstitial%3Furl%3Dhttps%3A%2F%2Fjournal.unsika.ac.id%2Findex.php%2Fsyntax%2Farticle%2Fdownload%2F227%2F223&usg=AOvVa>

- Yunis, R., & Telaumbanua, K. (2017). Pengembangan E-Learning Berbasiskan LMS untuk Sekolah, Studi Kasus SMA/SMK di Sumatera Utara. *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi (JNTETI)*, 6(1), 32–36. <https://doi.org/10.22146/jnteti.v6i1.291>
- Yustanti, I., & Novita, D. (2019). Pemanfaatan E-Learning Bagi Para Pendidik Di Era Digital 4 . 0. *Prosiding Seminar Nasional*, 338–346.
- Zareiforush, H., Minaei, S., Alizadeh, M. R., Banakar, A., & Samani, B. H. (2016). Design, development and performance evaluation of an automatic control system for rice whitening machine based on computer vision and fuzzy logic. *Computers and Electronics in Agriculture*, 124, 14–22. <https://doi.org/10.1016/j.compag.2016.01.024>