

Analyzing Student’s Learning Interests in the Implementation of Blended Learning Using Data Mining

<https://doi.org/10.3991/ijoe.v16i11.16453>

Yuri Ariyanto ^(✉), Budi Harijanto, Atiqah Nurul Asri
State Polytechnic of Malang, East Java, Indonesia
yuri@polinema.ac.id

Abstract—Blended Learning combines teaching and learning activities in the classroom and online teaching. In its implementation, this learning method requires a lot of data. One of them is the student's online test score data which can be used as an evaluation of learning. For this reason, in this study, data mining is used to determine the results of online student examinations as well as to determine student interest in learning about the implementation of Blended Learning. Data mining techniques are used in the logs of online learning session results, so that one can assess the online learning system used. By assessing the system, it can be identified which students who have studied hard and those who have not studied in the online exam. The series data used are student test score data on the State Polytechnic of Malang Learning Management System (LMS). The student score dataset is arranged based on variables in the Educational Process Mining (EPM) Dataset of UCI, which are obtained from teacher's assignments. In addition, data mining classification is used to determine student interest in learning towards blended learning. In the process of data mining, comparative analysis is carried out using the features of the RapidMiner tool to carry out the process of student data for training and data validation. This process uses several algorithms along with student data which is divided into two sets for training and validation. From the results of data mining calculations produce a classification with minimum errors. From the test, the resulting linear regression algorithm has RMSE 0.000 and SE 0.000, while the neural network algorithm has RMSE 0.525 and SE 0.275.

Keywords—Blended Learning, Data Mining, Classification Technique, Educational Process Mining, RapidMiner.

1 Introduction

In this era of higher education, the role of blended learning has provided considerable benefits [1][2][3]. Blended learning activities are mixed activities, which are a combination of activities from e-learning learning with face-to-face learning in class [4]. In general, the purpose of the blended learning initiative is to provide assistance that enables lecturers to help the instructors themselves in the teaching and learning process, with the aim of increasing students' understanding of face-to-face teaching and learning in class [4][5].

Quality resources depend on the role of decent higher education

[2]. Improving the quality of education must be done by improving the quality of learning, especially in terms of the processes and methods of learning carried out in higher education [6]. The concept of e-learning provides benefits for students to be able to access course material, information about academic operations, and online discussion forums between instructors and students [2] [7][8].

To process and analyze the education system, data mining techniques is used [9][10]. Data mining can help the educational institution in evaluating the education system policy, and in finding students' learning interests towards the learning system [10].

In this paper, the data mining classification method is applied to process and analyze data on IT Department student test scores in the Learning Management System (LMS) of State Polytechnic of Malang. The student score dataset variable used refers to the Educational Process Mining (EPM) Dataset of UCI and the processing and analysis of the score datasets uses the RapidMiner tool. Rapidminer processes data training datasets and data validation. The results of the study produce calculations from the data mining process, these results also reflect students' interest in the implementation of blended learning in teaching and learning activities.

2 Literature Review

2.1 Blended learning

Blended learning method of learning in general is a combination of learning. This method combines or integrates various types of methods, equipment, or used learning technology into face-to-face learning procedures in the classroom [1][2][3][4]. Blended learning is an innovation in learning methods where the learning and teaching process is carried out in new ways, while still referring to the rules of higher education [11][12][13].

2.2 Data mining

Data mining is widely used to increase knowledge in the world of education [14][15]. Data mining is an artificial intelligence tool that can be used to find useful information. Data mining processes data analysis from various dimensions and can categorize information [16]. In the process of analysis the DM algorithm can be used properly to provide solutions to achieve the desired results [10]. There are several techniques used in data mining such as classification, category, estimation and visualization [17][18]. The use of algorithms in data mining to do mathematical and specific calculations.

2.3 Rapidminer

RapidMiner is an interactive tool for conducting machine learning and data mining processes in implementing existing algorithms. Rapidminer software is open source, where free projects can be used by users free [19][20].

3 Research Method

The first stage of the research methodology is carried out by collecting data on student interactions on the use of blended learning with LMS State Polytechnic of Malang. The purpose of data collection is to provide information on the use of online learning systems with LMS. The LMS system provides information about student interactions in activities developed in the course, including student interactions with course topic files, assignments and evaluations for sending assignments, reading course material content, sending assignments and activities to teachers and online assessments. From these activities students are trained and given the opportunity to strengthen learning through content and assessments that exist on the LMS system.

The collected data is then analyzed and designed. This process produces a data set model that is used in testing. The data set consists of training data and validation data where the data variables are adjusted according to the UCI dataset EPM.

In general the testing phase plan consists of testing the quality and validation of the results obtained by the model. Data mining algorithm testing is done by selecting students' data in the form of test scores by dividing into 2 groups, one for training data and one as validation data.

The student score data set for training is 70% and the remaining dataset is used for validation so that it gives 100% of the data used for modeling. Based on the type of data used in the form of integers and real numbers in the dataset, the test is done by calculating the student's score estimation [17][21].

- Linear Regression uses 70% of data in accordance with the training dataset (T) and 30% of data is used for validation (V).
- Neural Network uses 70% of data in accordance with the training dataset (T) and 30% of data is used for validation (V).

Table 1. shows a description of the percentage calculation of students' final scores on the LMS. Student scores are obtained from the sum of 4 exercise sessions.

Table 1. Student Evaluation Dataset Description

Name	Description	Excercise Point
Student ID	Id of Student	-
ES 1.1	Exercise Session 1	20 %
ES 2.1	Exercise Session 2	20 %
ES 3.1	Exercise Session 3	30 %
ES 4.1	Exercise Session 4	30 %
Total	The grades total are given per student Id	100 %

4 Results and Discussion

To find out the interaction of students with the LMS, testing the student score dataset of 326 data is performed. Table 2. shows the performance testing of each algorithm using the RapidMiner tool.

Table 2. Algorithm Performance Results

Algorithm	Dataset	Examples Data	RootMeanSquaredError (RMSE)	SquaredError (SE)
Linear Regression	T 0.7	228	0.000	0.000
	V 0.3	98		
Neural Network	T 0.7	228	0.525	0.275
	V 0.3	98		

The graph of linear regression algorithm testing graph, for estimating students' scores on the use of the LMS system is shown in Fig. 1

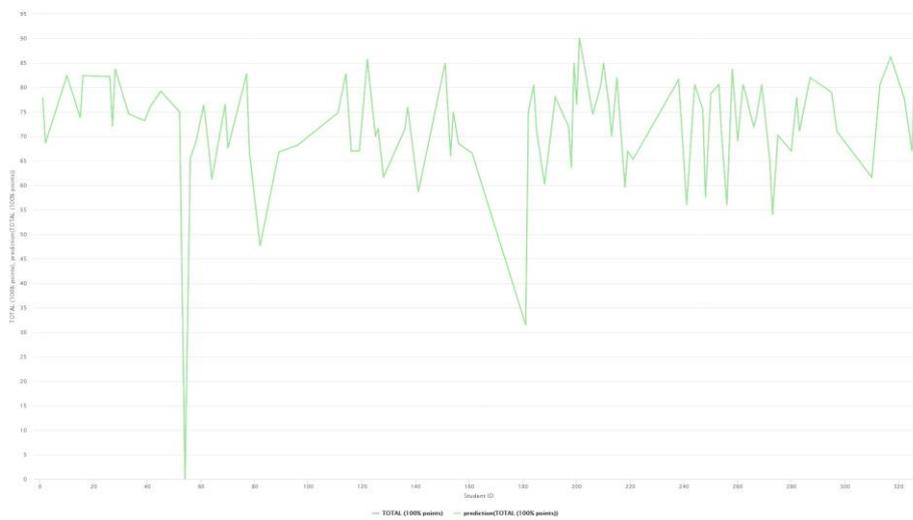


Fig. 1. Liner Regression Algorithm

In the Fig. 1. the total scores of students with the total predicted scores has the same value. The results of the liner regression algorithm process to predict the scores are shown in Fig. 2.

Row No.	TOTAL(100% points)	prediction(TOTAL(100% points))	Student ID
1	78	78.000	1
2	68.600	68.600	2
3	82.400	82.400	10
4	73.800	73.800	15
5	82.400	82.400	16
6	82.200	82.200	26
7	72	72.000	27
8	83.800	83.800	28
9	74.600	74.600	33
10	73.200	73.200	39

Fig. 2. RapidMiner Proses Liner Regression

A graph of a neural network algorithm testing graph for estimating student scores on the use of the LMS system is shown in Fig. 3.

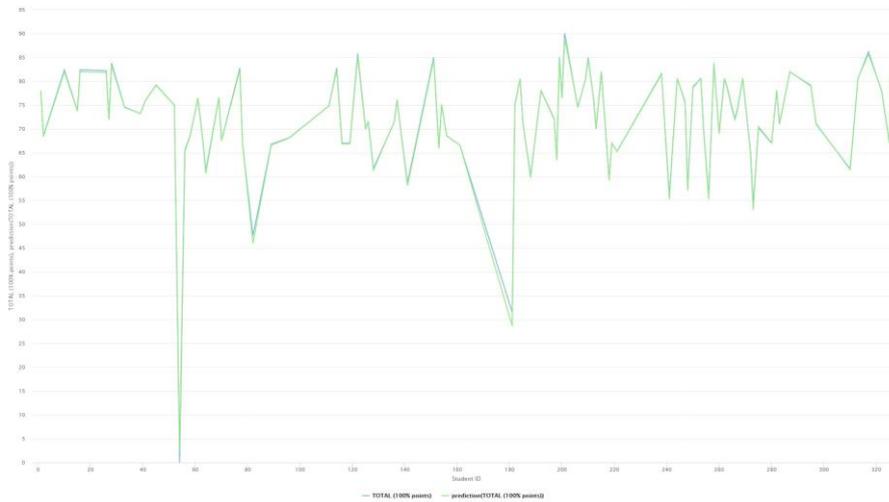


Fig. 3. Neural Network Algorithm

In the Fig. 3. the total score of students with the total of the prediction score differs in values. The results of the neural network algorithm process to predict scores are shown in Fig. 4., with predictive difference scores on student id 2, 10, 16, 26, 28 and 33.

Row No.	TOTAL(100...	prediction(T...	Student ID
1	78	77.961	1
2	68.600	68.396	2
3	82.400	81.985	10
4	73.800	73.762	15
5	82.400	81.989	16
6	82.200	81.847	26
7	72	71.987	27
8	83.800	83.311	28
9	74.600	74.479	33
10	73.200	73.194	39

Fig. 4. RapidMiner Proses Neural Network

5 Conclusion

From the results of testing the data mining algorithm, out of a total of 326 students, it can be concluded that the linear regression algorithm is better. Performance of linear regression algorithm with round and real value data types has an RMSE value of 0,000. With the appropriate algorithm performance calculation, it can be seen the value of student interaction with the blended learning system in terms of the use of the LMS system.

6 References

- [1] T. A. Weerasinghe, “An evaluation of different types of blended learning activities in higher education,” Proc. - IEEE 18th Int. Conf. Adv. Learn. Technol. ICALT 2018, pp. 42–45, 2018, <https://doi.org/10.1109/icalt.2018.00017>.
- [2] R. Sari and M. Karsen, “An empirical study on blended learning to improve quality of learning in higher education,” Proc. 2016 Int. Conf. Inf. Manag. Tech-nol. ICIMTech 2016, no. November, pp. 235–240, 2017, <https://doi.org/10.1109/icimtech.2016.7930336>.
- [3] C. Rui, “Research of blending learning making use of Web2.0 tools,” 2009 WASE Int. Conf. Inf. Eng. ICIE 2009, vol. 2, pp. 466–469, 2009, <https://doi.org/10.1109/ICIE.2009.253>
- [4] G. Frankl, “Evaluation of blended learning courses,” 2012.

- [5] T. Winterstein, F. Greiner, H. F. Schlaak, and L. Pullich, “A blended-learning concept for basic lectures in electrical engineering: A practical report,” 2012 Int. Conf. Educ. e-Learning Innov. ICEELI 2012, pp. 1–4, 2012, <https://doi.org/10.1109/iceeli.2012.6360578>.
- [6] S. Tikadar, S. Bhattacharya, and V. Tamarapalli, “A blended learning plat-form to improve teaching-learning experience,” Proc. - IEEE 18th Int. Conf. Adv. Learn. Technol. ICALT 2018, pp. 87–89, 2018, <https://doi.org/10.1109/icalt.2018.00027>.
- [7] B. Yi et al., “Learning analytics-based evaluation mode for blended learning and its applications,” Proc. - 2017 Int. Symp. Educ. Technol. ISET 2017, pp. 147–149, 2017, <https://doi.org/10.1109/iset.2017.42>.
- [8] S. M. Alyami and A. M. Alagab, “The difference in learning strategies in virtual learning environment and their effect on academic achievement and learning satisfaction for distance teaching and training program students,” Proc. - 2013 4th Int. Conf. e-Learning Best Pract. Manag. Des. Dev. e-Courses Stand. Excell. Creat. ECONF 2013, pp. 102–112, 2013, <https://doi.org/10.1109/econf.2013.40>.
- [9] A. Vilorio et al., “Determinating student interactions in a virtual learning environment using data mining,” Procedia Comput. Sci., vol. 155, no. 2018, pp. 587–592, 2019, doi: 10.1016/j.procs.2019.08.082.
- [10] S. Arumugam, A. Kovalan, and A. E. Narayanan, “A Study of Easy Educational Data Mining for E-Learning Log Data from complex and large Dataset,” vol. 11, no. 1, pp. 39–47, 2018.
- [11] D. Lalima and K. Lata Dangwal, “Blended Learning: An Innovative Approach,” Univers. J. Educ. Res., vol. 5, no. 1, pp. 129–136, 2017, <https://doi.org/10.13189/ujer.2017.050116>.
- [12] R. Saul, W. Wade, K. Hodgkinson, A. Smith, and J. Arfield, “Flexible Learning in Higher Education,” Br. J. Educ. Stud., vol. 44, no. 1, p. 129, 1996, doi: 10.2307/3121722.
- [13] W. D. Dwiyoogo and C. L. Radjah, “Effectiveness, efficiency and instruction appeal of blended learning model,” Int. J. online Biomed. Eng., vol. 16, no. 4, pp. 91–108, 2020, <https://doi.org/10.3991/ijoe.v16i04.13389>.
- [14] C. dos S. Garcia et al., “Process mining techniques and applications – A systematic mapping study,” Expert Syst. Appl., vol. 133, pp. 260–295, 2019, doi: 10.1016/j.eswa.2019.05.003.
- [15] X. X. Suhirman, J. M. Zain, and T. Herawan, “Data mining for education decision support: A review,” Int. J. Emerg. Technol. Learn., vol. 9, no. 6, pp. 4–19, 2014, <https://doi.org/10.3991/ijet.v9i6.3950>.
- [16] W. S. Ng, “Web Data Mining in Education,” vol. 7, no. 6, pp. 58–77, 2016, doi: 10.4018/978-1-5225-1877-8.ch005.
- [17] J. Cheng, “Data-Mining Research in Education,” no. March, 2017.
- [18] K. Sunday, P. Ocheja, S. Hussain, S. S. Oyelere, B. O. Samson, and F. J. Agbo, “Analyzing Student Performance in Programming Education Using Classification Techniques,” Int. J. Emerg. Technol. Learn., vol. 15, no. 02, p. 127, 2020, <https://doi.org/10.3991/ijet.v15i02.11527>.
- [19] A. Naik and L. Samant, “Correlation Review of Classification Algorithm Using Data Mining Tool: WEKA, Rapidminer, Tanagra, Orange and Knime,” Procedia Comput. Sci., vol. 85, no. Cms, pp. 662–668, 2016, <https://doi.org/10.1016/j.procs.2016.05.251>.
- [20] A. S. Hashmi and T. Ahmad, “Big Data Mining: Tools & Algorithms,” Int. J. Recent Contrib. from Eng. Sci. IT, vol. 4, no. 1, p. 36, 2016, <https://doi.org/10.3991/ijes.v4i1.5350>.
- [21] R. Hossain, S. M. H. Mahmud, M. A. Hossin, S. R. Haider Noori, and H. Jahan, “PRMT: Predicting Risk Factor of Obesity among Middle-Aged People Using Data Mining Techniques,” Procedia Comput. Sci., vol. 132, pp. 1068–1076, 2018, <https://doi.org/10.1016/j.procs.2018.05.022>

7 Authors

Yuri Ariyanto, is a senior lecturer at State Polytechnic of Malang, Indonesia and member of Institute of Electrical and Electronics Engineers. Interest in research is network technology, cloud computing, security system, Information Systems, and Software Engineering. Email: yuri@polinema.ac.id

Budi Harijanto, is an associate professor at State at Polytechnic of Malang, Indonesia. Interest in research is Artificial Intelligence, Information Systems, and Software Engineering. Email: budi.harijanto@polinema.ac.id

Atiqah Nurul Asri, is a senior lecturer at State Polytechnic of Malang, Indonesia. Interest in research is Educational Technology, Language Education, Vocational Education. Email: atiqah.nurul@polinema.ac.id

Article submitted 2020-06-21. Resubmitted 2020-07-09. Final acceptance 2020-07-09. Final version published as submitted by the authors.