■ 1223

Predicting the Presence of Learning Motivation in Electronic Learning: A New Rules to Predict

Christina Juliane*, Arry A. Arman, Husni S. Sastramihardja, Iping Supriana
School of Electrical Engineering and Informatics, Institut Teknologi Bandung (ITB),
Bandung, Indonesia
*Corresponding author, e-mail: christina.juliane@students.itb.ac.id

Abstract

Research affirms that electronic learning (e-learning) has a great deal of advantages to learning process, it's provides learners with flexibility in terms of time, place, and pace. That easiness was not be a guarantee of a good learning outcomes though they got the same learning material and course, because the fact is the outcomes was not the same from one to another and even not satisfy enough. One of prerequisite to the successful of e-learning process is the presence of learning motivation and it was not easy to identify. We propose a novel model to predicting the presence of learning motivation in e-learning using those attributes that have been identified in previous research. This model has been built using WEKA toolkit by comparing fourteen algorithms for Tree Classifier and ten-fold cross validation testing methods to process 3.200 of data sets. The best accuracy reached at 91.1% and identified four parameters to predict the presence of learning motivation in e-learning, and aim to assist teachers in identify whether student needs motivation or does not. This study also confirmed that Tree Classifier still has the best accuracy to predict and classify academic performance, it was reached average 90.48% for fourteen algorithm for accuracy value.

Keywords: Learning Motivation, Predicting, Tree Classifier

1. Introduction

Information and Communication Technology (ICT) and variety of approaches in the field of computer science have been used to provide easiness in the teaching and learning process. That easiness was not be a guarantee of a good learning outcomes though they got the same learning material and course, because the fact is the outcomes was not the same from one to another and even not satisfy enough. Many of e-learning models has been built to address these issues, including [1] build AVIEW (Amrita Virtual Interactive E-Learning World) which provides four types of real time collaborative multimedia which has ratings 95% for user friendly level. Reference [2] used the technology of Natural Language Processing (NLP), Machine Learning, Knowledge Representation and Ontology to build SESLATA (Semantic Self Learning and Teaching Agent) which intended to help students build knowledge in the form of conclusion from a teaching document. Reference [3] built Open Source-Based Mobile Learning that utilizes mobile technology such as smart phone to facilitate the distance issues of learning process, and [4] adapting facial recognition technology to build 3D Tutorial System to detect students emotions while they are learning.

Instead of mentioning the positive impact of e-learning to the process and learning objectives [1-4], the utilization of AVIEW motivated by the lecturers competencies that considered better than the lecturers in the traditional class [1]. As stated in [5], one of requirements must be fulfilled to be successful in e-learning is learning motivation, it has been proven that only certain people are motivated and successfully achieved the learning objectives in e-learning [6]. Motivation can be identified as a process that carried out by someone with intensity values, persistence, and effort direction to a goal achievement [7]. Measuring learning motivation in electronic learning is a process purposely to achieve certain learning objectives. This process is taken to build a conducive atmosphere in electronic learning activity. It is initially by identifying the attributes of measurement that can be used. Reference [8] has identified three groups of attributes used to measure a learning motivation in electronic learning, including velocity (v1, v2 and v3), quantity (q) and relevancy (r1 and r2) of the answers that have been

given by students in electronic learning process. But those attributes have not be a real solution to measure a person's learning motivation.

This study being extend from previous research [8] that already identified a new parameters to measured learning motivation and also propose a new model to predicting the presence of learning motivation in e-learning. The model had clarified the parameters to measured learning motivation that conducted in [8] and propose a new one. The model was built by Tree Classification Technique, it is because a research in the field of classification and prediction in the context of academic performance ratings has previously been carried out using Tree Classifier and Non Tree Classifier [10], [13-14]. The Tree Classifier algorithm is used to build the prediction model of learning motivation, considering its simplicity to interpret the results by giving a number of rules of making decision or prediction [9]. Another reason is the level of accuracy that can be achieved above 95% [10]. The study then showed that Tree Classifier algorithm such as C5.0, J48, Simple Chart, Rep Tree, NB Tree, CART, QUEST, or CRUISE had a better accuracy for classification, compared to Non Tree Classifier algorithms such as NN, Naïve Bayes, Bayes Net, INN, and MLN [10-14]. The organization of the paper is as follows. Part 2 is to present the methodology used in the study and Part 3 is result and discussion, that explained how to build a prediction model in the assessment of learning motivation. The final conclusion and acknowledgment of the study are discussed in Section 4 and Section 5 respectively.

2. Research Method

The research methodology consists of four stages (Figure 1). The first stage is the process of data preparation to prepare a list of questions and select the potential respondent related to the field of the study. The second phase is data collection, which was conducted by distributing the questionnaires to identify the size of learning motivation using the attributes from previous research [8]. The third stage refers to the process of data processing, conducted by data mining techniques in WEKA. The fourth stage is data testing, it was conducted by comparing the accuracy of Tree Classifier Algorithm in WEKA by 10-fold cross validation technique until it successfully identified the best accuracy. A prediction model from the best algorithm with the highest level of accuracy would be used as the rules to measure the presence of learning motivation.

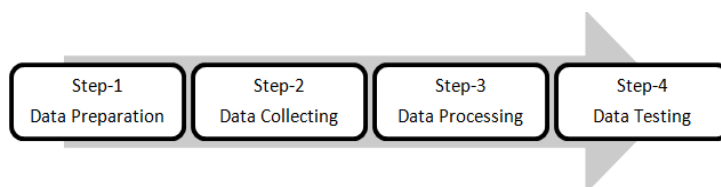| Step-1 Data Preparation	Step-2 Data Collecting	Step-3 Data Processing	Step-4 Data Testing

Figure 1. Research Methodology

3. Results and Discussion

The modeling process of predicting the learning motivation in electronic learning consisted of four parts, which are the process of data mining tool selection, preprocessing of data collection, modeling, and model evaluation.

3.1. Data Mining Tool Selection

Tree Classifier toolkit for WEKA was selected to process the data in consideration to that WEKA is an open source tool [15] that is simply to be used [12] and to be applied in data mining matters and based upon GUI (Graphical User Interface) [10], Tree classifier algorithms are used in data training process. The rules that have been produced by the best algorithm would then be selected to measure the learning motivation.

3.2. Data Collection and Preprocessing

Data collecting process in this study was produced by distributing the questionnaires that consisted of list of questions to identify the size of learning motivation using the attributes from [8]. A total of 3.200 data sets have been collected from 533 of selected potential respondents. The respondents selected from students and lectures between age 20 to 50 years old and taking place in Java province-Indonesia that related to the field of the study. The list of questions was generated from the analysis of secondary data from [8] and identify the value of it by spreading it online. The value options inspired from true and false logic [10], so it would be easy to understand and giving the value. The attributes that we are used in the classification as seen in Table 1.

Table 1. Attributes Candidate

Group of Attributes	Symbol	Description	Value
Velocity	v_1	Speed of answering questions/assignments/quizzes/exams/email/chat	Fast-Slow
	v_2	Speed of downloading learning material/assignment/ exams/quizzes question	Fast-Slow
	v_3	Speed of uploading assignments answer/ quizzes answer/ tests answer	Fast-Slow
Quantity	q	Quantity (a length) of the answer of assignment/exams/quizzes	A Bit-Many
Relevancy	r_1	The correctness of answering questions/assignment/exams/email	False-True
	r_2	Following the rules or standards	Neglect-Follow
Motivation	m	The value of learning motivation	Low-High

3.3. Modeling

A prediction model was built by classifying the collected data sets. 10-fold cross validation was selected for data testing, meaning that 3.200 numbers of instances would be split into 10 parts, each of which would be randomly formed by the principle of a 10-fold validation with 1: 9. It means that one part would be used as data testing, and other nine would be used as data training. The process was continued until all ten parts had an opportunity to be a data testing and the accuracy was measured [9]. The training process was performed by distributing 3.200 numbers of instances and divided it into five groups [16] of random sample of data sets. The data distributing consisted of 640, 1280, 1920, 2560, and 3.200 random sample of data sets and used to comparing the accuracy of tested algorithm.

3.4. Model Evaluation

Data testing was divided into the multiple sets of instances [16] each of which was tested with 14 types of Tree Classifier algorithm in the WEKA toolkit to see the test result of each algorithm. The result indicators we used here consisted of time used in building the model (time/second), accuracy, precision, recall, f-measure, and ROC (Receiver Operating Characteristic) [17]. The test results then indicated that the Random Tree, ID3, and FT algorithm became three algorithms with the best test result in each group of instances (Table 2). However, Even FT algorithm had the best value for accuracy, the ROC value is less than ID3 and Random Tree. And considered the time taken to build the model, ID3 was slower than Random Tree. So, it meant that the best test result was indicated from Random Tree. This means that the rule generated from the best algorithm would be used as a prediction model to measure learning motivation, namely the rule resulted from Random Tree Algorithm for the amount of data in 1280.

The accuracy result of Random Tree Algorithm for multiple distributed data sets is presented in Table 3 showing that the Random Tree algorithm for the numbers of 1280 instances was more advantageous in terms of time used in building the model (time/second), accuracy, precision, recall, f-measure, and ROC (Receiver Operating Characteristic) categories compared with other numbers. There were, however, shortcomings in terms of ROC values that were smaller when compared to the numbers of 1920 instances and it took a quite a long time for modeling when being compared to other four groups of instances. Despite having a smaller value of ROC, the diagnosis of ROC was still in the good category of classification [18], and the

time required to build the model was still quite fast at 0.24 seconds when compared with the time taken on ID3 algorithm on the same numbers of instance. The comparisons chart of precision, recall, and F-measure for Random Tree algorithm in numbers of instances is presented in Figure 2.

Table 2. Algorithm Accuracy for 1280 of instances

No	Type of Tree	Time/s	Accuracy	Precision	Recall	F-Measure	ROC
1	AD Tree	0.16	90.00	0.89	0.90	0.89	0.86
2	BF Tree	2.00	91.00	0.90	0.91	0.90	0.86
3	Decision Stump	0.01	87.03	0.76	0.87	0.81	0.69
4	FT	1.13	91.25	0.90	0.91	0.90	**0.84**
5	Id3	**0.24**	91.10	0.90	0.91	0.90	**0.86**
6	J48	0.01	90.50	0.89	0.90	0.90	0.82
7	J48graft	0.08	90.50	0.89	0.90	0.90	0.82
8	LAD Tree	0.37	90.90	0.90	0.90	0.90	0.86
9	LMT	2.79	91.00	0.90	0.90	0.90	0.88
10	NB Tree	1.57	90.10	0.89	0.90	0.89	0.88
11	Random Forest	0.08	90.90	0.90	0.90	0.90	0.86
12	Random Tree	**0.00**	91.10	0.90	0.91	0.90	**0.86**
13	REP Tree	0.07	90.70	0.89	0.90	0.90	0.82
14	Simple Cart	1.23	90.70	0.89	0.90	0.90	0.82

The percentage comparison for accuracy value for Random Tree algorithm had the best value on the numbers of 1280 instances and it can be seen in Figure 3. The prediction model produced by Random Tree algorithm would be used as the rules to measure the learning motivation for electronic learning. The rules that have been obtained had 41 numbers of classification rules, 36 of those numbers were to predict a high learning motivation and 5 rules were to predict a low learning motivation. A total of 41 rules that have been generated can be grouped into 11 groups of the new rules by seeing the similarity patterns of classification. The new classification group consisted of 6 rules to predict a high learning motivation and 5 rules to predict a low learning motivation. Table 4 presents the new group classification rules.

Table 3. Test Result of Random Tree Algorithm for five distributed data sets sample

Instance	Correctly	Incorrectly	Precision	Recall	F-Measure	ROC	Time/s
640	87.80	12.20	0.86	0.88	0.86	0.81	0.03
1280	91.10	8.90	**0.90**	**0.91**	**0.90**	0.86	0.00
1920	89.90	10.00	0.89	0.89	0.89	0.88	0.00
2560	86.60	13.40	0.79	0.86	0.80	0.77	0.01
3200	84.59	15.40	0.79	0.85	0.80	0.79	0.01

The rules for the identification of high learning motivation consisted of six rules and had four main attributes consisting of r1, v3, r2, and q attributes. Otherwise, the value of r1, v3, r2, and q attributes when being predicted to a low learning motivation had a zero value (0) on all generated rules. A special characteristic identified on rule number 10 is presented in Table 4, when attributes of r1, v3, r2, and q had zero value, the value of v1 attribute and v2 also had the same value (0). The value of v1 and v2 attributes was not a major cause of a low learning motivation prediction, so those attributes were being ignored because the major valuation was on the attributes of r1, v3, r2, and q. Therefore, it can be concluded that to predicting the value of learning motivation was quite represented by r1, v3, r2, and q attributes. Hence, the new attributes used to measure learning motivation in the context of electronic learning had four attributes (Table 5), and the new rules that were used to measure learning motivation in electronic learning had 9 rules (Table 6).
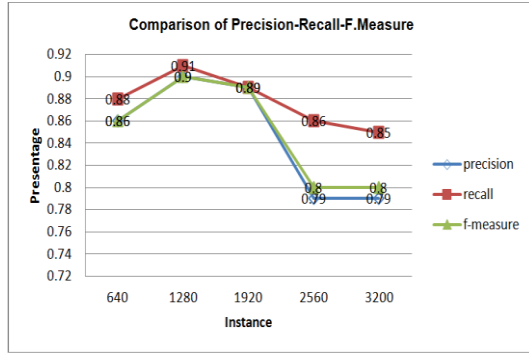
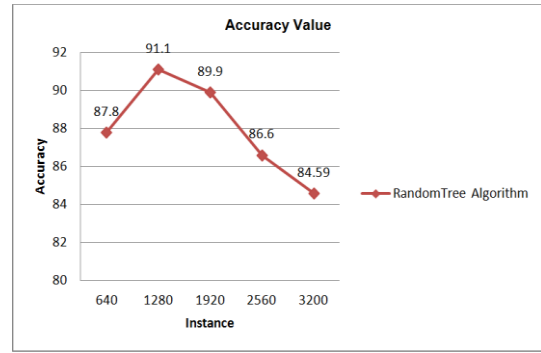Figure 2. Precision-Recall-F.Measure of RandomTree Algorithm for 5 groups of Instance

Figure 3. Accuracy Value of RandomTree Algorithm for 5 groups of Instance

Table 4. Rules from RandomTree Classifier to Predicting Learning Motivation for e-Learning

No	Type of Motivation	Rules
1		If the answer is correct, then the learning motivation is high. ($r1 = 1 \rightarrow m=1$
2		If the answer is wrong but fast in uploading the answer, then the learning motivation is high.($r1 = 0, v = 1 \rightarrow m = 1$)
3	High Learning Motivation	If the answer is wrong but the rules or standards are followed, the the learning motivation is high.($r1 = 0, r2 = 1 \rightarrow m = 1$)
4		If the answer is wrong and but the rules or standards are neglected, then the learning motivation is high.($r1 = 0, r2 = 0, q = 1 \rightarrow m = 1$)
5		If the answer is wrong, but its many, then the learning motivation is high.($r1 = 0, q = 1 \rightarrow m = 1$)
6		If the answer is wrong but rules or standards are followed, then the learning motivation is high.($r1 = 0, r2 = 1 \rightarrow m = 1$)
7		Although it is fast in uploading the answer and fast in answering too, but if the answer is a bit and wrong, do not follow the rules or standards, Then the learning motivation is low.($v1, v3 = 1 \,\&\, r1, r2, q = 0 \rightarrow m=0$
8		Although it is fast in answering, but If the answer is wrong, slow in uploading the answer, the answer is a bit and do not follow the rules or standards, then the learning motivation is low.($v1 = 1 \,\&\, v3, r1, r2, q = 0 \rightarrow m=0$
9	Low Learning Motivation	Although it is fast in downloading, but if the answer is wrong, slow in uploading and answering, and do not follow the rules or standards, then the learning motivation is low.($v2 = 1 \,\&\, v1, v3, r1, r2, q = 0 \rightarrow m = 0$
10		If it is slow in answering, downloading, uploading and the answer is a bit, wrong, neglect the rules or standards, the the learning motivation is low.($v1, v2, v3, q, r1, r2 = 0 \rightarrow m = 0$)
11		Although the answer are many and fast, but if it is wrong, slow in uploading, neglect the rules or standards, then the learning motivation is low.($q, v1 = 1 \,\&\, r1, r2, v3 = 0 \rightarrow m = 0$)

Table 5. New Attributes in Measurement of Learning Motivation on Electronic learning

Group of Attributes	Symbol	Description	Value
Velocity	v	Speed of uploading assignments answer/quizzes answer/ tests answer	Fast-Slow
Quantity	q	Quantity (a length) of the answer of assignment/exams/quizzes	A Bit-Many
Relevancy	r_1	The correctness of answering questions/assignment/exams/email	False-True
	r_2	Following the rules or standards	Neglect-Follow
Motivation	m	The value of learning motivation	Low-High

Table 6. New Rules to Measure Learning Motivation on Electronic Learning

No	Type of Motivation	Rules
1		$(r1 = 1 \rightarrow m = 1)$
2		$(r1 = 0, v = 1 \rightarrow m = 1)$
3		$(r1 = 0, r2 = 1 \rightarrow m = 1)$
4	High Learning Motivation	$(r1 = 0, r2 = 0, q = 1 \rightarrow m = 1)$
5		$(r1 = 0, q = 1 \rightarrow m = 1)$
6		$(r1 = 0, r2 = 1 \rightarrow m = 1)$
7		$(v = 1 \, \& \, r1, r2, q = 0 \rightarrow m = 0)$
8	Low Learning Motivation	$(v, r1, r2, q = 0 \rightarrow m = 0)$
9		$(q = 1 \, \& \, r1, r2, v3 = 0 \rightarrow m = 0)$

4. Conclusion and Future Work

The Tree Classifier algorithm is used to build the prediction model of learning motivation is simple to interpret. This study has identified that Random Tree algorithm has the best accuracy compared to another algorithm in WEKA Tree Classifier toolkit, it was reached 91,1%. This study also confirmed that Tree Classifier still has the best accuracy to predict and classify academic performance, it was reached average 90.48% for fourteen algorithm for accuracy value, even though it was not good enough compared to [10]. Random Tree algorithm identified 9 rules to measure learning motivation in the context of electronic learning, which are 6 rules to predicting a high learning motivation and 3 rules to predicting a low learning motivation (Table 6). Furthermore, elimination occurs for v1 attribute (the speed of answering questions/assignments/quizzes/exams/email/chat) and v2 attribute (the speed of downloading material/task/exam/quiz questions in [8]. It was ignored because the value of those attributes did not affect the value of learning motivation. Therefore, the results of this study can be identified that motivation is a function of the speed of uploading assignments answer/ quizzes answer/ tests answer (v), quantity (a length) of the answer of assignment /exams /quizzes (q), the correctness of answering questions/ assignment/ exams/ email (r1), and follows the rules or standards (r2).

$$m = f(v, q, r_1, r_2) \tag{1}$$

References

[1] Kamal B, Jayahari KR, Ancy M. AVIEW: Real Time Collaborative Multimedia e-Learning. *Journal of MTDL-ACM.* 2011.

[2] Koggalahewa DN, Amararachi, Pilapitiya, Gegange. *Semantic Self Learning and Teaching Agent (SESLATA).* The 8th International Conference on Computer Science & Education (ICCSE 2013). Colombo-Sri Lanka. 2013.

[3] Arief H, Victor GU. Open Source Based M-Learning Application for Supporting Distance Learning. *TELKOMNIKA Indonesian Journal of Electrical Engineering.* 2014; 12(3): 657-664.

[4] Ahmad Basori H, Andi T, Andi BFM. Intelligent Avatar on E-learning using Facial Expression and Haptic. *TELKOMNIKA Indonesian Journal of Electrical Engineering.* 2011; 9(1): 115-124.

[5] Leung CH, Yuen YC. Knowledge management System for Electronic Learning of IT Skills (SIGITE'07). *Journal on ACM.* 2007.

[6] Bart R, Dirk T, Piet vd B, Wim G, Mien S. The Role of Academic Motivation in Computer-Supported Collaborative Learning. *Journal of Elsevier 2009 in Human Behavior.*

[7] Sandra G, Bernard W. Theories and Principles of Motivation. Los Angeles: University of California.1996.

[8] Christina J, Arry AA, Husni SS, Iping S. *Measurement of Learning motivation in Electronic Learning.* International Conference on Information Technology Systems and Innovation (ICITSI). Bandung-Bali, Indonesia. 2015.

[9] Vercellis C. Business Intelligent-Data Mining and Optimization for Decision Making. New York: John Wiley & Sons, Ltd. 2009.

[10] Taruna S, Mrinal P. An *Empirical Analysis of Classification Techniques for Predicting Academic*

Performance. IEEE International Advance Computer Conference-IACC. 2014.

[11] WEKA. University of Waikato. New Zealand.

[12] Luan. Data Mining Application in Higher Education. SPSS. Inc. 2004.

[13] Diego GS, Marta Z. *Comparing Classification methods for Predicting Distance Students Performance*. JMLR. Workshop and Conference Proceedings. 2011.

[14] Pandey M, VK Sharma. A Decision Tree Algorithm Pertaining to the Student Performance Analysis and Prediction. *International Journal of Computer Application*. 2013; 61(13).

[15] Bidgoli BM, Kashy DA, Kortemeyer G, Punch W. *Predicting Student Performance: An Application of Data Mining Methods with an Educational Web-based Systems*. Proceeding of 33rd Frontiers in Education Conference. 2003: 13-18.

[16] Jesse D, Mark G. *The Relationship Between Precision-Recall and ROC Curves*. Proceeding of the 23rd International Conference on Machine Learning. Pittsburgh. 2006.

[17] Powers DMW. Evaluation: From Precision, Recall and F-Measure to ROC,. *Journal of machine Learning Technologies*. 2011; 2(1): 37-63.

[18] Raghavan, Ratheesh. Study of the relationship of training set size to error rate in yet Another Decision Tree and Random Forest Algorithms. Thesis. Computer Science at Texas Tech University; 2006.