ISSN (Print) : 0974-6846
Indian Journal of Science and Technology, Vol 9(14), DOI: 10.17485/ijst/2016/v9i14/89029, April 2016 ISSN (Online) : 0974-5645

A Systematic Approach to Identify Unmotivated Learners in Online Learning

P. V. Praveen Sundar* and A. V. Senthil Kumar

Hindusthan College of Arts and Science, Coimbatore – 641028, Tamil Nadu, India; praveensundarpv@gmail.com, avsenthilkumar@yahoo.co.in

Abstract

Objectives: To make an online learning to be effective, it is necessary to identify the unmotivated learners and motivate them to avoid attrition. **Methods/Analysis:** Through this paper, we identify the unmotivated learners using log file analysis. Usually in a log file analysis, time spent on learning alone is not enough to determine the motivational level of the learners, because some learners may understand the concepts very quickly, some may take more time to understand the concepts. Hence it is difficult to conclude the motivational level of the learners using time spent attribute alone. **Findings:** In our educational system, the student is qualified based on the marks they secured in an exam. The marks only decide whether he is engaged or disengaged in a study. Thus our proposed model will identify the unmotivated learners based on the learning time along with the traditional assessment marks. This improves the prediction performance of unmotivated learners and it becomes very compatible in online learning. **Improvements:** We compare and discuss our results with traditional log file based approaches. The results show that our proposed methodologies will give better results than the traditional log file based approaches.

Keywords: Disengagement Detection, Enhanced Disengagement Detection Algorithm (EDDA), Learning and Assessment based Methodology, Log File Analysis, Online Learning

1. Introduction

Learning is often considered to be an essential part in a human life. Nowadays learning is considered as tool for getting job and some may have considered it as a tool for achieving knowledge. Whatever the goal of learning it is unavoidable at any circumstances. Learning came across so many stages and nowadays, most of the universities, colleges, institutions prefer online learning and it is considered as a great alternative for traditional class room learning. Through online learning, learners have a chance to study at their flexible place and time, the learner may participate in several online courses all around the world to increase their skills. Over the advantages, the main disadvantage of the online learning is to identifying the motivational level of the learner. It is difficult to identify whether the learner is motivated or unmotivated and identifying the reason for unmotivated is also difficult in online learning. Whereas in traditional learning, the instructor will identify it using behaviour, attendance, body language etc. and offer interventional strategies aimed to increase motivation. Hence we need a mechanism to monitor the learner's activities and find out the reason for their unmotivated, which results an online learning to be more effective.

Motivation is recognized as an important contributor to learning. "Motivation can influence what, when and how we learn"[1]. "Many instructors consider the motivation level of learners is an important factor in successful instruction"[2]. Online learners are not like traditional students in their need of encouragement but they need encouragement and support at the right time and the right place to keep them on their path to learning[3]. Motivating a learner can take many forms as[4] describes that having the clear goals, concentrating and focusing on goals, balance between ability level and challenge leads the learner

**Author for correspondence*

to be engaged[5]. States that engagement of the learner will depend on positive and negative thoughts, persistence, effort, goals and self-regulated behaviours. The learner's hesitancy in performing actions after reading the tutorial, based on the task, that the learner will complete in a specific period[6] and how frequently the learners requested for help to complete the tasks will inform the tutor to infer the learner confidence and also if the learner is searching external content for a related topic it may be a sign of getting lost in the course content; it may also be a sign of an elaboration cognitive strategy[7,8]. Research helps the instructors to view online learning platforms from the student's perspective and arranges the information based on the good architecture principles that enable users to find data with a few mouse clicks. Farzan et al. proposes a Course Agent system[9], which helps students to select courses that are most relevant to their career goals. Debbie Morrison[3] mentions that the responding the learners queries quickly will motivate the learners. Even though the learners seem unmotivated or poor in learning, giving a constructive and supportive feedback will help us the learner to be engaged. [10]Proposes an initial groundwork on detecting behavioural disengagement via quitting behaviour it operationally defines a point at which a student opts to stop interacting with the learning activity.

2. Disengagement Detection Methodology

Time spent is considered to be an important factor of detecting the disengaged learners. According to[11] examines various attributes like time spent on reading, number of mouse click, time spent on moving/scrolling the mouse etc. and concludes that time spent on reading a page is an important indicator for finding disengagement behaviour of a learner. [12]States that the fast moving on pages, as well as students who spend long pauses are more likely to learn just shallowly. They did not gain the knowledge of the learning material and most probably they seem to be disengaged. The learner's engagement can be identified based on the average session duration and time on task percentage[13]. According to[14] engagement is determined based on the two metrics, i.e. too short time to read texts and to answer questions or taking too long time read or answering questions. According to[15] analyzed the behaviour of the student in terms of average time spent in online and category of visited websites by them along with their academic performance.

2.1 Redefining Threshold Values

The researchers[16–20] agrees that 5 seconds has their minimum threshold but they have different maximum thresholds for detecting disengaged learners. Maximum threshold of 600 seconds and they make a footprint of mouse click on every page and calculated the TSR value for deducting disengagement in[17]. The research of[16,19] has a maximum threshold of 420 seconds. If 2/3 sequences lie between 5 and 420 seconds means then the user is considered as engaged, otherwise the learner will be considered as disengaged. [20]Also considered that 5 seconds has minimum threshold and the learner with no activity for more than 420 seconds in a page (which includes scrolling mouse and clicking the mouse on sub links in that page) is also considered as disengagement.

The threshold values calculated by the[16,17,19,20] researchers are based on the time spent values of their dataset, hence there is a need of finding new threshold values for our dataset. The time interval for reading and number of pages in each interval for our dataset is presented in Table 1.

Table 1 indicates that most of the pages require less than 240 seconds to read a page, similarly 6182 which means less than 1% of pages requires more than 720 seconds. Hence we assign the range for finding minimum threshold is less than 240 seconds and maximum threshold is greater than 720.

The formula for finding minimum threshold value is:

$$\sum_{i=m}^{n} ai = a_m + a_{m+1} + a_{m+2} + a_{m+3} + \ldots\ldots a_{n-1} + a_n \tag{1}$$

$$\mu = \frac{1}{n}\sum_{i=1}^{n} ai \tag{2}$$

For minimum threshold, where i = 1, n = Total number of pages read on the given threshold value, and a = total time spent for the given minimum threshold value.

Table 1. Time interval for number of pages read in each interval

Time Interval	No of page's read
< 240 seconds	4,76,864
>240 and <480 seconds	2,08,380
>480 and <720 Seconds	21,554
>720 Seconds	6182

Similarly, for maximum threshold, where i = 720, n = Total number of pages read on the given threshold value and a is the total time spent of a page on given maximum threshold values.

$$\text{Minimum threshold } (\mu 1) = \frac{1}{476864} * 76,67,974 = 16.08 \quad (3)$$

$$\text{Maximum threshold } (\mu 2) = \frac{1}{6182} * 53,44033 = 864.45 \quad (4)$$

$$\begin{aligned}\text{Exact Pages Read} = {} & \text{Total no of pages read} \\ & - (\text{Total no of pages above threshold} \\ & + \text{Total no of pages below threshold})\end{aligned} \quad (5)$$

2.2 Log File Analysis

As mentioned earlier, the proposed work identifies the unmotivated learners through log file analysis. The main advantage of log file analysis is without interrupting the learner, the instructor can observe the learner's activities. Once the learner logged into the system, their activities are stored in a separate log file. Thus all learners are belonging to the focused group, thereby information about the learners are well informed.

The log file entries are classified into three groups namely learning, assessment and other sequences. Assessment sequences are grouped as individual tests based on the login and logout. The overall values are stored in database. Using other sequences, learner's attitude, effort and interest is calculated. Using Enhanced Disengagement Detection Algorithm, learning sequences are monitored and it identifies the motivational level of the learners. Thus the scope of this proposal is not limited to predict the disengagement alone. This proposal can act as an aid to explore various problems of the learner.

2.3 Enhanced Disengagement Detection Algorithm (EDDA)

Enhanced Disengagement Detection Algorithm is used to construct and predict the disengagement based on redefining the threshold values on learning and marks they scored in assessment. Through log file analysis, each and every learning sequence is monitored. If the learning sequence is less than minimum threshold value, then the sequence is assigned as disengaged. Similarly, if the sequence is greater than the maximum threshold means then it has to check further condition that, whether there are any activities happened on those time spent (which includes mouse activities). If any activities happen on those sequences means then the system will have considered that sequence as slow learner and assigns that sequence as Engaged or else the sequence is assigned as Disengaged. Once the learning sequences are monitored, the system has to find the Exact Pages Read by the learner and then the system checks if 2/3 of the total number of pages read is greater than the Exact Pages Read (EPR) and the learner has to get at least 50% of correct answers in their assessment means then the system will assign the Overall status of the learner as Engaged learner, otherwise the learner is considered as a Disengaged learner.

Enhanced Disengagement Detection Algorithm (EDDA)

Initialize log file sequences $lf_1, lf_2, lf_3 \ldots\ldots lf_n$
Output: Preprocessed log file with engagement status.
Step 1: Begin.
Step 2: For each sequences in log file lf_i do.
Step 3: Assign Status = 'Disengaged'.
Step 4: If time spent< µ1 then.
Step 5: Go to Step 2.
Step 6: Else if time spent> µ2 and NoS = 0 and NoMC = 0 then.
Step 7: Go to Step 2.
Step 8: Else assign Status = 'Engaged'.
Step 9: End If.
Step 10: End For.
Step 11: Calculate EPR using Equ (5).
Step 12: For each item in Student Database do.
Step 13: If ((EPR < (2/3 * NoP)) and (Noc> .= NoQ/2)).
Step 14: Assign Eng_status = 'Disengaged'.
Step 15: Else Assign Eng_status ='Engaged'.
Step 16: End If.
Step 17: End For.
Step 18: End.

3. Experimental Results and Discussion

3.1 Dataset Preparations

In order to validate our approach, we have collected the log files of 247 users from an online learning system namely Quasi framework[21], where each learner has spent minimum of ten sessions for learning and ten sessions for exam activities. The proper login and logout is considered as session. Totally 7,90,859 instances have been obtained. Out of those instances, 7,12,980 instances are identified as learning instances and 49,623 instances is identified as assessment instances and other activities like feedback, glossary, getting help has occurred 28,256 instances. Totally 33 attributes are derived from logged events. The list of logged events is presented in[20]. The Hybrid PSO with Naïve Bayes classifier is used for feature selection[22]. After feature selection process, the selected attributes used for this analysis are listed in Table 2.

In addition to the above attributes, we add three meta attributes to our research, 1) NoAT : Number of pages above threshold established for maximum time required to read a page (865 seconds), 2) NoBT : Number of pages below threshold established for minimum time required to read a page (17 seconds), 3) EPR: Exact Pages Read by the learner. Those attributes are considered as a meta attributes because they are derived from the raw data.

This research detects the disengagement based on the new threshold values and compares the results with previous approaches. We use two datasets for this analysis, DS_1 includes six attributes related to reading pages and taking assessments and DS_2 includes eight attributes related to reading pages, taking assessments and Mouse dynamic attributes. Using these two datasets we have obtained four result sets namely F1, F2, F3, F4. In F1 result set is generated based on[19] approach, where the minimum threshold is fixed to 5 seconds and maximum threshold is fixed to 420 seconds. F2 result set is generated based on the[20] approach, where minimum threshold is fixed to 5 seconds and maximum threshold is fixed with 420 seconds and check the mouse event attributes. F3 result set is generated based on our proposed methodology, where minimum threshold is calculated based on the Equation (3) and maximum threshold is calculated based on the Equation (4). F4 result set generated using the same methodology followed by F3 dataset with an extra condition of Mouse related events.

Table 2. Attributes used for analysis

Code	Attributes Description
NoP	No of Pages Read
AvgTL	Average Time Spent for Learning
NoQ	Number of Questions Attended
AvgTQ	Average time spent on Assessment
NoC	Number of Correct Answers
NoW	Number of Wrong Answers
NoMC	Number of Mouse clicks used
NoS	Scrolls wheels used

3.2 Performance Analysis

True positive and accuracy is considered as a key factor to confirm the quality of our proposed methodology, similarly other indicators such as false positive rate, precision, error rate are also calculated. The output of Four result sets is displayed in Table 3:

Among the experimental results we obtained, it confirms that the new threshold values will give better results than the previous approaches.

3.3 Confusion Matrix

The confusion matrix of F1, F2, F3 and F4 result set is presented in Tables 4(a), 4(b), 4(c) and 4(d).

The fact that there is a small variation between the F3 and F4 result set. The best performance indicates that

Table 3. Experimental results

	F1	F2	F3	F4
% Correct	89.07	92.71	93.12	93.93
True Positive Rate	0.932	0.940	0.944	0.926
False Positive Rate	0.246	0.109	0.087	0.045
Precision	0.927	0.961	0.938	0.962
Error	0.109	0.073	0.069	0.061

Table 4(a). Confusion matrix of F1 result set

		Predicted		
		Disengaged	Engaged	Total
Actual	Disengaged	177	13	190
	Engaged	14	43	57
	Total	191	56	247

Table 4(b). Confusion matrix of F2 result set

		Predicted		
		Disengaged	Engaged	Total
Actual	Disengaged	172	11	183
	Engaged	7	57	64
	Total	179	68	247

Table 4(c). Confusion matrix of F3 result set

		Predicted		
		Disengaged	Engaged	Total
Actual	Disengaged	136	8	144
	Engaged	9	94	103
	Total	145	102	247

Table 4(d). Confusion matrix of F4 result set

		Predicted		
		Disengaged	Engaged	Total
Actual	Disengaged	125	10	135
	Engaged	5	107	112
	Total	130	117	247

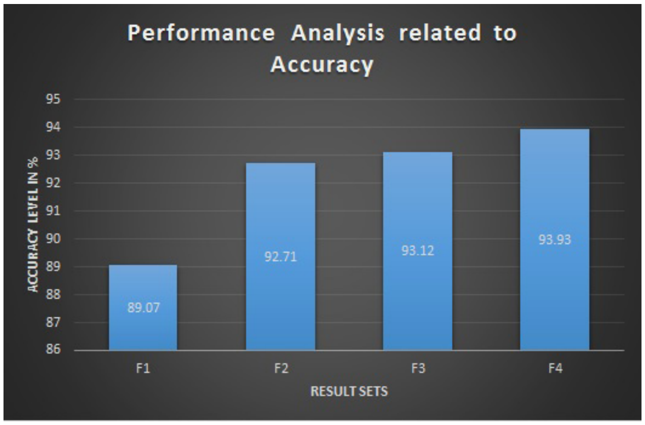

Figure 1. Performance analysis related to accuracy.

attributes related to the mouse dynamics with new threshold values will produce the high prediction values. While adding the mouse events with new threshold values, it is identified that 4 learners are identified as slow learners and they are moved to Engaged status and two Engaged learners who are wrongly classified as Disengaged is also rightly classified as Engaged. When comparing the F3 result set with F1 and F2 result sets, it confirms that the most of the engaged learners are considered as Disengaged, due to the improper thresholding setting to the current dataset. Thus the new threshold values will predict the Disengaged learners accurately and produces the higher accuracy level than the other result sets.

3.4 Chart Values

Figure 1 shows that the fixed threshold values will produce lesser prediction values and redefining new threshold

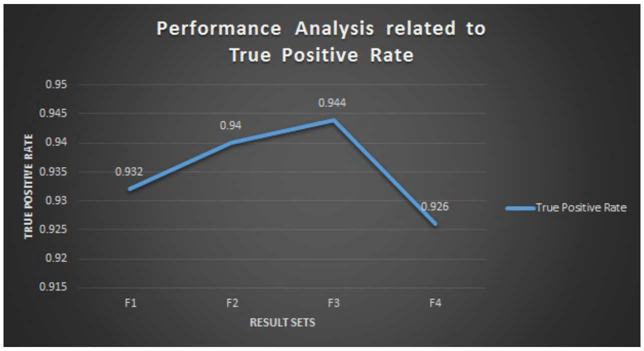

Figure 2. Performance analysis related to true positive rate.

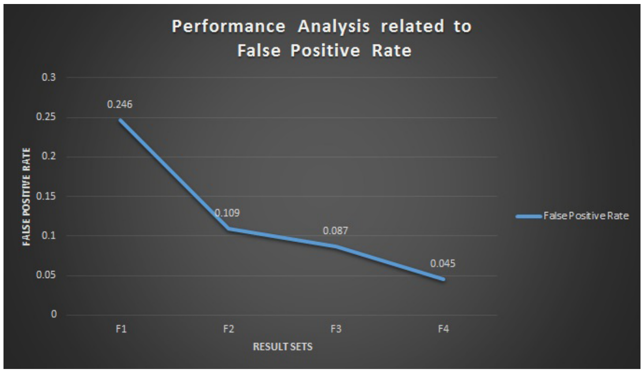

Figure 3. Performance analysis related to false positive rate.

values will increase the accuracy level. Similarly adding Mouse attributes to the result sets will slightly increase the accuracy level.

Figure 2 shows that true positive rate is highly increased in proposed result sets (F3, F4).

Figure 3 shows that false positive rate is highly decreased in proposed result sets.

4. Conclusion

Disengagement detection is considered to be an essential factor in online learning systems and earlier detection of disengaged learners will make an online learning system to be a successful one. As seen in review of literature, disengagement detection is predicted using so many methodologies. Most of the systems find out the disengaged learners using time spent attribute. In our proposed system, disengaged detection is predicted using learning and assessment based methodology. In a learning based prediction, time spent attribute is considered as a key factor for detecting the Disengaged learners, the new threshold values based on the time spent on page will give better prediction values than the previous approaches. As learning alone cannot be enough for finding the Disengaged learners, thus we include their assessment results and conclude their engagement status. Compared to the previous proposals, it has unified structure to redefine the disengagement prediction logic, but this can be further explored in many aspects to confirm the consistency and reliability of the Quasi framework.

5. References

1. Schunk D, Meece J, Pintrich P. Motivation in education: Theory, research and applications. 4th ed. Pearson; 2014.
2. Dick W, Carey L, Carey J. The systematic design of instruction. 7th ed. Pearson; 2009.
3. Morrison D. How to motivate students in the online learning environment. 2012. Available from: https://onlinelearninginsights.wordpress.com/2012/08/31/how-to-motivate-students-in-the-online-learning-environment/
4. Mihaly C. Finding Flow: The psychology of engagement with everyday life. New York: Basic Books; 1st ed. 1997.
5. Fredricks J, McColskey W, Meli J, Montrosse B, Mordica J, Mooney K. Measuring student engagement in upper elementary through high school: A description of 21 instruments. Issues and Answers Report; REL. 2011 Jan. 98(098). p, 88.
6. Qu L, Johnson WL. Detecting the learner's motivational states in an interactive learning environment. Proceeding of the 2005 Conference on Artificial Intelligence in Education; Amsterdam, Netherlands. 2005. p. 547–54.
7. Pintrich P, Schunk D. Motivation in Education: Theory, Research and Applications. 2nd ed. Englewood Cliffs, NJ: Prentice-Hall; 2002.
8. Gordesky J, Marmon M. Designing online platforms to increase motivation for learning. Proceedings of University of Texas College of Information. 2013 Research Exchange Conference; 2013. p. 60–2.
9. Farzan R, Brusilovsky P. Encouraging user participation in a course recommender system: An impact on user behaviour. Computers in Human Behaviour. 2011 Jan: 27(1):276–84.
10. Mills C, Bosch N, Graesser A, Mello SD. To quit or not to quit: Predicting future behavioural disengagement from reading patterns. Stefan Trausan-Matu, Kristy Elizabeth Boyer, Martha Crosby and Kitty Panourgia, editors. Intelligent Tutoring Systems. Lecture Notes in Computer Science. Springer International Publishing. 2014; 8474:19–28.
11. Claypool M, Le P, Waseda M, Brown D. Implicit interest indicators. Proc International Conference on Intelligent User Interfaces; Santa Fe. 2001. p. 33–40.
12. Gowda SM, Baker RS, Corbett AT, Rossi LM. Towards automatically detecting whether student learning is shallow. International Journal of Artificial Intelligence in Education. 2013 Nov; 23(1):50–70. Doi No.: 10.1007/s40593-013-0006-4.
13. Hershkovitz A, Nachmias R. Developing a log-based motivation measuring tool. Baker RSJD, Barnes T, Beck JE (Eds.). Educational Data Mining 1st International Conference on Educational Data Mining; Montreal, Quebec, Canada. 2008 Jun 20-21. p. 226–33.
14. Monterrat B, Desmarais M, Lavoue E, George S. A player model for adaptive gamification in learning environments. AIED; Madrid, Spain. 2015; 9112:297–306.
15. Oskouei RJ, Askari M, Sajja PRP. Differential impact of web technology on various dimensions of students' academic and non-academic activities - A case study. Indian Journal of Science and Tehnology. 2013 Jul; 6(7):4912–22.
16. Cocea M. Assessment of motivation in online learning environments. M.Sc. National College of Ireland; 2007.
17. Farzan R, Brusilovsky P. Social navigation support in E-Learning: What are real footprints. Proceedings of IJCAI'05 Workshop on Intelligent Techniques for Web Personalization; Edinburgh, U.K. 2005. p. 49–56.
18. Sundar PVP. Quasi Framework: A new student disengagement detection in online learning. International Journal of Engineering Research and Technology (IJERT). 2012 Dec; 1(10):1–6.
19. Sundar PVP, Kumar AVS. Disengagement detection in online learning using Quasi framework. IOSR Journal of Engineering (IOSRJEN). 2015; 5(1):4–9.
20. Sundar PVP, Kumar AVS. An enhanced disengagement detection in online learning using Quasi framework. Special Edition of International Journal of Applied Engineering Research (IJAER). 2015 Jun; 10(55):1298–302.
21. Quasi Framework. Available from: www.quasiframework.com
22. Sundar PVP, Kumar AVS. A hybrid classification method for disengagement detection in online learning. International Journal of Education and Information Studies. 2015; 5(1):67–74.