IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
Vol. 05, Issue 01 (January. 2015), ||V1|| PP 04-09

Disengagement Detection in Online Learning Using Quasi Framework

P.V.Praveen Sundar[1], A.V.Senthil Kumar[2]
[1]*Research scholar, Hindusthan College of Arts and Science,Coimbatore.*
[2] *Director of MCA, Hindusthan College of Arts and Science, Coimbatore.*

Abstract: **-** Online Learning enables to deliver both learning and information dynamically. It allows tapping the knowledge of experts and non-experts and catapulting those messages beyond classroom walls and into the workplace but it is more difficult to assess a learner's motivation. In traditional classroom learning, when a student shows signs of de-motivation, teachers know how to motivate their students and how to optimize their instruction. But in On-line learning it is difficult to measure their disengaged level. By tracing the disengaged learners in online learning, will help us to motivate the learners at proper time. Thus making an online learning to be a successful one. Thus we propose a new framework called Quasi Framework, which is trying to measure the significant relationship between disengagement level and their academic achievement. In this paper we compare our proposed framework with ihelp, a web based learning system.

Keywords: *Disengagement Detection, Online Learning, EDM, Log File Analysis.*

I. INTRODUCTION

On-line learning is a new emerging technology which is dynamic and potentially enriching forms of learning but attrition remains a serious problem [1]. Motivation towards learning is affected by the learner's self efficacy, locus of control, goal orientation and perceived task difficulty. In a traditional classroom environment, tutors infer learners' levels of motivation from several cues, including speech, behavior, attendance, body language or feedback, and offer interventional strategies aimed at increasing motivation. Online Learning System needs an ability to recognize when the learner is becoming de-motivated and to intervene with effective motivational strategies. Such a system would comprise two main components, an assessment mechanism that infers the learner's level of motivation from observing the learner's behavior, and an adaptation component that selects the most appropriate intervention strategy to increase motivation.

To address this challenge, we restricted our research to one motivational aspect, disengagement, and looked at identifying the relevant information from learners' actions to be used for its prediction. Being able to automatically detect disengaged learners would offer the opportunity to make online learning more efficient, enabling tutors and systems to target disengaged learners, to reengage them, and thus, to reduce attrition. The motivational based disengagement detection system is still in infancy stage with respect to assessing the learner's attitude and characteristics. Thus we propose a new framework for disengagement detection to enhance the value of existing prediction systems.

The learner's actions preserved in log files have been recently discovered as a valuable source of information and several approaches to motivation detection and intervention have used log-file analysis. An important advantage of log-file analysis over self-assessment approaches is the unobtrusiveness of the assessment process, similar to the classroom situation where a teacher observes that a learner is not motivated without interrupting his/her activities.

Analyzing data from log-file is an efficient method for automatic analysis, whereas it has certain level of fuzziness in order to retrieve desired information in robust fashion. Thus, we introduce metadata as quasi assessment technique to obtain the result. Further, disengagement is correlated with academic achievement to ensure the quality of assessment.

Among different studies accomplished on disengagement detection, through this paper we compare our proposed work with Cocea et al. [2] study. Cocea et al.[2] had conducted two validation studies on iHelp data indicate that the attributes identified in the studies on HTML-Tutor data are relevant for the new system as well. Paired t-tests were used to investigate the statistical significance of the differences in the distribution of accuracy and true positive rates across the eight methods between the two studies on iHelp data on the other hand, the mean for each data set and the significance of the t-test are considered. All accuracy and True-Positive (TP) rates on all data sets were tested and proved to follow a normal distribution.

When comparing the results of two iHelp studies, we can see that the difference is statistically significant with one exception, i.e., the difference between the accuracy distribution for the data sets with sequences of only 10 minutes (DS1_S1 and DS1_S2). The amount of data and the new score attribute did not contribute to better predictions.

To assess the contribution to prediction of the attributes in each system, three attribute evaluation methods with ranking as search method for attribute selection were used: chi-square, information gain, and OneR [12]. The attribute ranking using information gain filter for iHelp attributes delivered the following ranking: NoPpP, NoPages, AvgTimeP, NoPpM, AvgTimeQ, Score, and NoQuestions. Chi-square evaluator produces the same ranking, except that the positions of the last two attributes are reversed, i.e., NoQuestions contributes a higher gain than Score. OneR evaluator produces a different ranking compared to the other two, even if the main trend is preserved (attributes related to reading come before the ones for quizzes): NoPpP, AvgTimeP, NoPages, NoPpM, NoQuestions, AvgTimeQ, and Score.

The attribute ranking results shows that the attributes related to reading are more important than the ones related to tests. This study suggested that despite the problem they may pose, knowledge about the two patterns of disengagement would be useful for a more targeted intervention and in further work; the possibility to predict them will be investigated. The next section discuss about our new framework proposed for this study.

II. QUASI FRAMEWORK

Quasi Framework- is an online learning and assessment based environment [9], having three subjects namely Data Structures and algorithms, Introduction to computers and Java Programming. An entire course material is written in English. In the screenshot displayed in Figure-1, the topics are listed in left side of the screen. The contents of the selected topics are displayed in the center part of the page. The menu bar contains links to Assessment environment, frequently asked questions, preferences, glossary, and help on how to use the system, statistical tool about the personal usage of the system, search, remarks and feedback. The page navigation is displayed in bottom of the page.

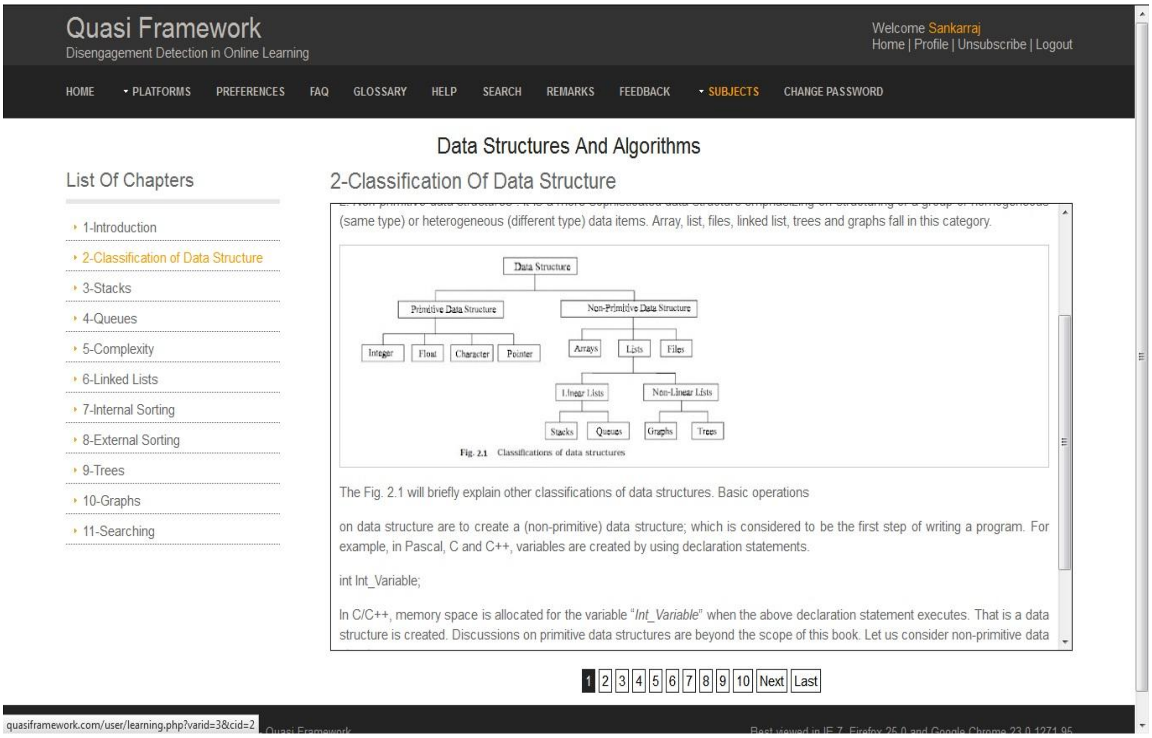

Fig. 1: Screenshot of Quasi Framework from Data Structures and Algorithms Subject.

The subsequent Figure-2 exhibits the structure of Quasi framework, which contains log information and database oriented information about the students. In the first phase, information fetched from the respective files and it is directed to pre-processing which do the basic pre-processing procedures. After pre-processing it is directed to merging section, which merges the log information and database information. From the source of information provided for prediction module, training is suggested for to construct the new model. Further, it predicts the disengagement based on the common rules followed by [2][10][11].

Cocea et al. study[2] will be very suitable for predicting generic disengagement prediction, whereas for disengagement to engagement model should consider continuous evaluation method. Hence predicting the personalized disengagement pattern helps to cluster the students.

A generic student disengagement prediction will not suitable for all scenarios, since ability and characteristics of study may differ from one another. Some students have the ability to understand the concept within short period of time and perform well in their examinations and some of them may not follow the same structure.

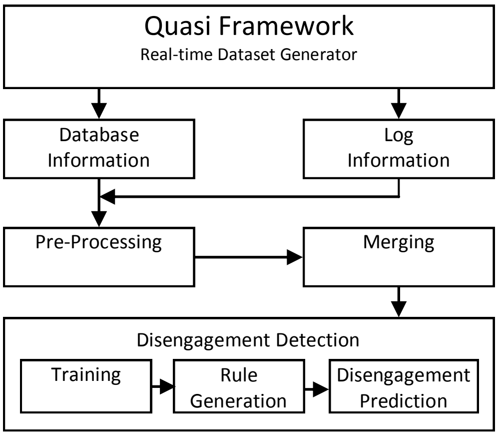

Fig. 2: Structure of Quasi Framework

As per the general online learning environment student enroll it in a course. Further online tutoring option will be provided to learn the specific subject and our proposed work generates log file, which observe all the required attributes such as time spent on specific concept, page, navigation pattern, number of times logged in to a system. Similarly Assessment option contains only online examination towards the specific subject proposed for him. The entire activity should be written in log-file. Quasi framework is suggested to use the meta-data format for student observations. Prior academic history will be separately maintained for effective assessment.

III. EXPERIMENTAL RESULTS

For this experiment, we have considered the real time dataset, which is collected from [9]. There are 25 students participated in our research dataset and each user has spent minimum 20 learning and assessment sessions, where every session is marked by login and logout. From this process 21,237 instances has been obtained. From the logged events, a total of 32 attributes were derived. Two events— reading pages and taking tests— occurred considerably more often than all the others, with a frequency of occurrence of 14,063 and 5585, respectively, out of a total of 21,237 sequences. Two other events— Checking their performance and statistical tool— were noticeably more frequent than the rest, with a frequency of 198 and 229, respectively, while the remaining events were rare (with an average of 23 occurrences in 21,237 sequences).

As [10][11] suggests that if a learner has spent less than 5 seconds or more than 420 seconds(7 minutes) on a page means, he/she will be labeled as disengaged for that particular sequence. If the 2/3 of the sequences are between 5 and 420 seconds, then the learner is considered as engaged, if the 2/3 of the sequences are not between 5 and 420 seconds means, then the learner is considered as disengaged learner.

To check the consistency of our prediction, we check the results with following eight classification algorithms of WEKA tool [12].

1) Naïve bayes updateable (NB)
2) Bayes Net (BN)
3) Naïve bayes Tree (NBT)
4) Random Forest(RF)
5) Random Tree(RT)
6) J48 Graft(JG)
7) One-R (1R)
8) Alternative Decision Tree (ADT).

In the analysis, we have taken six attributes related to reading pages and taking assessments. These attributes are presented in Table-1. We use two datasets for this analysis, DS_1 that includes all the sequences and DS_2 that includes all the sequences but limited to 10 minutes sequences.

Table 1: Attributes used for analysis

Code	Attribute Description
NoP	No of Pages read
AvgTL	Average Time spent for learning
NoQ	No of Questions attended
AvgTQ	Average time spent on assessment
NoC	Number of Correct Answers
NoW	Number of Wrong Answers

To check the accuracy level of our prediction values, two indicators are mainly considered: Accuracy and True positive rate, where accuracy is calculated as the percentage of correctly classified instances divided by total number of instances. True positive for disengaged class is an indication of the correct identification of disengaged learners, similarly other indicators such as false positive rate, precision, error are calculated.

For DS_1 and DS_3, 70 percent of the sequences were used for training and 30 percent were used for testing and for DS_2 and DS_4, 60 percent of the sequences were used for training and 40 percent were used for testing. The results of the above dataset are displayed in Table-2.

Table 2: Experimental results of Quasi Framework

Dataset	Results	NB	BN	NBT	RF	RT	JG	1R	ADT
DS_1	**% Correct**	80	80	84	84	84	84	84	84
	TP Rate	0.67	0.89	0.84	0.75	0.67	0.67	0.78	0.79
	FP Rate	0.13	0.25	0.16	0.1	0.06	0.06	0.13	0
	Precision	0.75	0.67	0.84	0.83	0.86	0.88	0.78	1
	Error	0.2	0.43	0.21	0.2	0.19	0.26	0.15	0.17
DS_2	**% Correct**	88	84	86	88	88	84	84	84
	TP Rate	1	0.9	0.67	0.78	0.89	0.89	1	1
	FP Rate	0.5	0.33	0.06	0.06	0.13	0.19	0.25	0.67
	Precision	0.86	0.9	0.86	0.88	0.8	0.73	0.69	0.83
	Error	0.21	0.16	0.34	0.18	0.23	0.17	0.45	0.16

The highest percentage of correctly predicted instances was obtained by naïve bayes updateable classification. The confusion matrix for naïve bayes updateable classification is presented in Table-3. Focusing of disengaged learners only, out of 8, three of the classification produce true positive rate as 1.

Table 3: Confusion matrix for naïve Bayes updateable

		Predicted		
		Disengaged	**Engaged**	**Total**
Actual	**Disengaged**	8	1	9
	Engaged	2	14	16
	Total	10	15	25

When compared to the Cocea et al [2] study, the accuracy level of DS_1 (including all sequences) lies between 76.51 and 83.13 but our proposed method, the accuracy level of DS_1 (including all sequences) lies between correct prediction of classified instances lie between 80 and 84. Similarly DS_2(limited to 10 Min Sequences), The accuracy level of Cocea et al[2] study lies between 81.90 and 85.34 but our proposed method, the accuracy level of DS_2(limited to 10 Min Sequences) lies between 84 and 88. Likewise True positive rate of DS_1 of Cocea et al Study lies between 0.64 and 0.82 but our proposed method true positive rate lies between 0.67 and 0.9. Similarly True positive rate of DS_2 of Cocea et al Study lies between 0.90 and 1 but our proposed method true positive rate lies between 0.67 and 1.

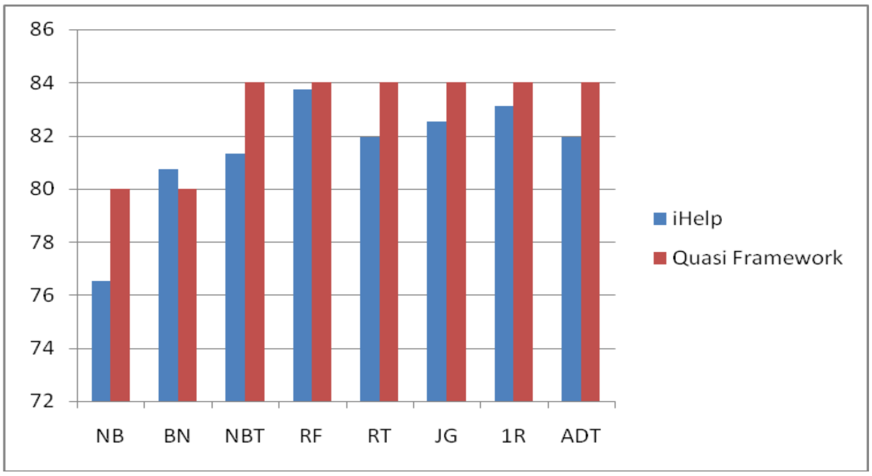

Fig. 3: Accuracy level of iHelp and Quasi framework for All Sequences

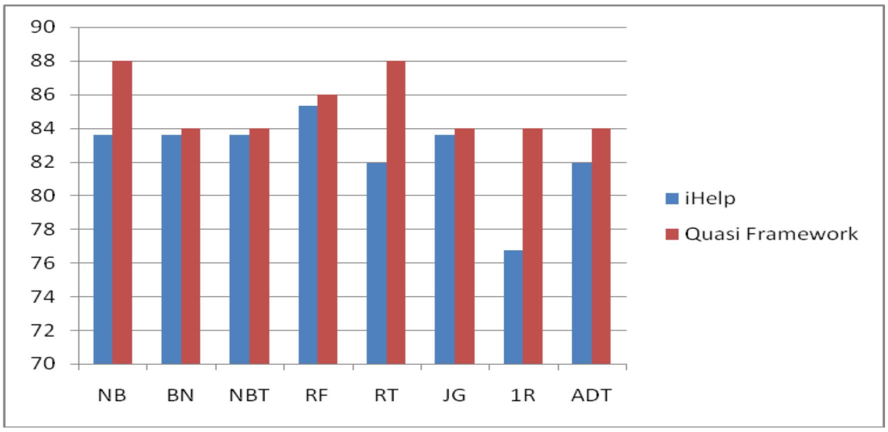

Fig. 4: Accuracy Level of iHelp and Quasi Framework for 10 Min Sequences

Figure 3 shows the Accuracy level of ihelp and Quasi framework for all sequences and Figure 4 shows the accuracy level of ihelp and Quasi Framework for 10 min sequences. The above result shows that Quasi Framework has a better accuracy level than ihelp dataset.

IV. CONCLUSION

Disengagement detection is a supportive mechanism to keep the students engaged in their academic activities. So many online tutoring models have been proposed to maximize the chance of engagement. Since online tutoring models are in growing interest, student engagement and their performance should be maintained as better than of traditional teaching method. Through this paper the comparison of two learning systems iHelp and Quasi framework has been done in detail. The comparative study had shown that quasi framework has good accuracy level than ihelp learning system. In iHelp reading is considered as a most valuable attribute than ones related to tests, but in Quasi framework, reading and assessment have equal importance on disengagement detection and we strongly accept that time spent alone cannot be very effective on disengagement detection, so adding new attributes will predict more disengaged students and give more accuracy.

REFERENCES

[1] Dille, B., & Mezack, M.: Identifying predictors of high risk among community college telecourse students. *The American Journal of Distance Education, (1991) 5(1), 24-35.*

[2] Mihaela Cocea and Stephan Weibelzahl “Disengagement Detection in Online Learning: Validation Studies and Perspectives”, *IEEE Transactions on Learning Technologies, Vol. 4, NO. 2, April-June 2011,Page 114-124.*

[3] Nussbaum & Steele, Situational disengagement and persistence in the face of adversity. *Journal of Experimental Social Psychology, Volume 43, Issue-1, 2007, Page 127–134.*

[4] Major & Schmader, “Coping With Ethnic Stereotypes in the Academic Domain: Perceived Injustice and Psychological Disengagement”, the Society for the Psychological Study of Social Issues, 1998.

[5] R'enger & Loose, "Relationship of sociocultural factors and academic self-esteem to school grades and school disengagement in North African French adolescents", *British Journal of Social Psychology, 45, 777–797. 2006.*

[6] Schmader et al, "Coping with ethnic stereotypes in the academic domain: Perceived injustice and psychological disengagement", *Journal of Social Issues, 57, 93–111. 2001.*

[7] Tougas, F., Rinfret, N., Beaton, A. M., & De la Sablonniere, R. "Policewomen acting in self defense: Can psychological disengagement protect self-esteem from the negative outcomes of relative deprivation", *Journal of Personality and Social Psychology, 88, 790–800. 2005.*

[8] P.V.Praveen Sundar, "Quasi Framework: A New Student Disengagement Detection in Online Learning", *International Journal of Engineering Research & Technology (IJERT), Vol. 1 Issue 10, December- 2012.*

[9] Quasi Framework, http://www.quasiframework.com, 2012

[10] Mihaela Cocea and Stephan Weibelzahl,"Cross-System Validation of Engagement Prediction from Log Files"," Creating New Learning Experiences on a Global Scale Lecture Notes in Computer Science Volume 4753", *Springer-Verlag Berlin Heidelberg 2007, pp 14-25*.

[11] Mihaela Cocea and Stephan Weibelzahl, "Validation issues in educational data mining–the case of HTML-Tutor and iHelp"," Handbook of Educational Data Mining-2010, pp 377-388.

[12] I.H. Witten and E. Frank, *Data Mining. Practical Machine Learning Tools and Techniques, second ed. Morgan Kauffman/Elsevier, 2005.*