Automated Discrimination of The Learners Attitude in Online Learning

P. V. Praveen Sundar[1]
Dr. A. V. Senthil Kumar[2]

Abstract:

Online learning environment is new platform in education arena, which provides handsome of features to the learners. Despite the online learning system seems attractive it is greatly affected by the poor monitoring capabilities. These issues had handled through behavioural and psychological facts by many researchers known as disengagement detection. The objective of the disengagement detection is to extract the causes and pattern of disengagements. Once the disengagement is identified, it is necessary to motivate the disengaged learners. The proposed work discriminates the learner's attitude using TSVM approach and extracts the causes and pattern of disengagements, which helps the instructor to motivate the disengaged users. Our experimental result states that the proposed method classifies the disengaged users well and makes the online learning system effective.

Keywords: *Online learning, Discrimination, Disengagement Detection, Learning attitudes*

I. INTRODUCTION

Online learning is the new learning mechanism which is considered as an internationally recognised qualification and the learners does not need to attend classes on campus. There are many terms used to describe learning that is delivered online, via the internet, ranging from Distance Education, to computerized electronic learning, online learning, internet learning and many others. We define online learning as courses that are specifically delivered via the internet to somewhere other than the classroom where the professor is teaching. It is not a course delivered via a DVD or CD-ROM, video tape or over a television channel. It is interactive in that you can also communicate with your teachers, professors or other students in your class. Sometimes it is delivered live, where you can "electronically" raise your hand and interact in real time and sometimes it is a lecture that has been pre-recorded. There is always a teacher or professor interacting /communicating with you and grading your participation, your assignments and your tests. Online learning has been proven to be a successful method of training and education.

The biggest issue that might arise in the online learning system is the lack of learner interests on the online learning system. Some of the factor of learning system might change the learner focus which would lead to failure of the system. Thus online learning systems should capable of finding the learners behaviours and changes to make them to focus on the assessment. This would be more difficult for some of online learner which needs to consider for the better performance [1].

Learning attitude is defined as a thoughts and actions of the individual person in doing some works. Every learner would have their own learning attitude which would differ with other learner attitudes. So, single evaluation method may not be suited for every learner. Each learner should treat with the different evaluation method based on their learning attitude. However, finding the learning attributes of the online learners with the concern of the differentiating their learning behaviour would be more difficult process.

The overall research of this work concentrates on discriminating the learning attributes of the online learners. The proposed research work automates the disengagement detection process by introducing the machine learning approach namely Transductive support vector machine (TSVM). This approach would learn the training data set that contains various online learners' opinion about the learning system. After finding the learner who is disengaged. The disengaged learners would be classified into two classes namely Partially disengaged and Fully disengaged. This classification is done using the naive bayes classification algorithm. After classification of the disengaged learners, the system assigns tasks based on learner's classification.

The organization of the proposed research work is given as follows: In the section 2, various previous research works that has been conducted in terms of finding the disengagement activities of users. In the section 3, detailed discussion of the proposed research methodology is given. In the section 4, experimental evaluation of the proposed research methodology is given in depth. In the section 5, Final conclusion of the research method is given.

II. RELATED WORKS

Research undertaken in the area of attitude and attitude formation shows that attitudes and beliefs are linked, and attitudes and behaviours are linked; moreover, attitudes are essentially divided into likes and dislikes [2]. With the broad expansion of ICT in education during the last decade, many research studies have explored the attitudes of users (educators and students) towards the integration of ICT in education [3][4][5]. Similarly many research studies analysis

[1]Research scholar, Hindusthan College of Arts and Science, Coimbatore, E-mail: praveensundarpv@gmail.com
[2]Director of MCA, Hindusthan College of Arts and Science, Coimbatore, E-mail: avsenthilkumar@yahoo.com

the learners attitudes and identifies the unmotivated learners in online learning [6][7][8][9][10][11].

University students in developing countries have varying attitudes towards e-learning but generally their attitudes are positive [12]. This was emphasised in [13] who pointed out that many students had positive attitudes towards e-learning because it had a positive impact on their motivation as well as self-esteem.

In some developing countries, the learning process and the adherence to traditional practices are inseparable. Thus, technology-based tools for e-learning are viewed as an interference with the practices that have been valued for generations. A good example of this scenario is Botswana. The socio-cultural environment in Botswana is very strong. Students in Botswana's higher learning institutions are still so strongly embedded in it that their attitude towards e-learning reflects it. Despite having taken significant steps towards a Western-style economy and towards urbanisation, the country maintains strong connections to its traditional roots. According to [14], students gain most of their knowledge from their integration with the various communities that have strong values, knowledge and beliefs.

However, the reserved attitude towards technology-based learning tools in Botswana and other developing countries is gradually disappearing, as the obvious usefulness of these tools is being realised. As Internet-enabled mobile devices are becoming more and more popular in developing countries, so is the range of areas in which they are applied; and this includes e-learning. Mobile devices such as Internet-enabled phones are very popular and are increasingly being used for blogging and social networking; this, in turn, helps improve user attitudes towards e-learning.

According to [15] however, the physical separation between the learner and the instructor tends to create a feeling of isolation on the part of the learner leading to negative attitudes.

A 2002 survey carried out in Pakistan's Virtual University with 387 undergraduate students in their final year of study concluded that over 90% of the students viewed learning through satellite TV and the Internet as advantageous, and student attitude towards e-learning were generally positive [16]. In [17] identified student attitudes as a factor that determined how e-learning was adopted in Iran. In [18] stated users who were very familiar with web technologies and the skills needed to use computer and mobile devices for instruction developed positive attitudes. On the other hand, students who were not skilled in ICT became anxious about the use of computers, had lower expectations from educational technology, and they did not believe in the benefits of e-learning [19].

III. DISCRIMINATING LEARNING ATTITUDES

Online learning is the most important and growing field in the real world environment which eliminates the need for students or learners to show up in the class room to learn the new things. Online learning system eliminates the maintenance of class room system with various technologies and professional. And also students can learn from anywhere without reaching the school. However, some drawback of online learning system might reduce the involvement online learners. This process of not involving in the online learning system is called as disengagement activity. This disengagement activity would occur due to many reasons like inconvenient environment, inconvenient learning time schedules, improper learning materials and soon.

Generally traditional evaluation system would follow the fixed pattern for all learners. This single and fixed traditional evaluation system may not be suitable for the real world learners who are having different learning attitudes. In this research work, learners are discriminated in terms of their disengagement activity that occurs because of the various learning attitudes. The overall research work is automated by introducing the machine learning approach namely Transductive support vector machine. This will train the log files of the various online learners based on which active users would identified as disengaged or not. After finding the disengaged learners those would classified into two classes to discriminate them. The classes that are used for discrimination are, "Partially disengaged and Fully disengaged". Based on these classification results, learners would be treated with different methodologies, so that disengagement possibility can be reduced considerably. The overall flow diagram is given in Figure- 1.

The processing step of proposed research methodology is listed as like follows:

- Learn the log files of online learners
- Find the disengaged learners among active users
- Classify the disengaged learners based on the learning attitude
- Assign tasks based on their class

These steps are explained clearly in the following sub sections.

Figure 1. Overall flow of Proposed research work

A. LEARN THE LOG FILES OF ONLINE LEARNERS

In the proposed research work, the log files of 120 learners are gathered from the online learning environment namely Quasi Framework [20]. The identified log files will be learned and categorized in accordance to their activity behaviour. This learning is done by using the machine learning algorithm called TSVM which aims to categorize the log files that are identified based on their type and nature. This TSVM algorithm is used to avoid the convergence problem which may occur due to the large volume of data.

TSVM is the iterative algorithm which aims to cover the unlabelled data present in search space with the help of transductive function by separating the hyper plane optimally. In the training phase of the TSVM, unlabelled data will be considered for the classification. That is the log files without label will be considered in the training phase for the better classification process. By using the TSVM methodology, better generalization capability of the classifier can be achieved. Also, this methodology aims to move the hyper plane gradually based on reaching a finer place. This methodology aims to achieve a finer place by achieving the generalization capability of the classifier. A finer hyper place position can be achieved by iteratively doing this mechanism.

In this approach, better classification can be achieved by migration of the data into the class labels which pretend to avoid the misclassification of the data by avoiding the discriminate functions. This mechanism also aims to provide more accurate results than the existing methodology. The convergence of the categorization depends on the similarity present among the different class labels.

ALGORITHM

Input: Labelled points: $S = [(x_j, y_j)]$, $j = 1, 2, \ldots, l$ and unlabelled points: $V = [(x_j)]$, $j = l+1, \ldots, n$.

Output: Transductive SVM classifier with original training set and the transductive set.

Begin

Initialize the working set $W^{(0)} = S$, previous transductive set $A^{(0)}{}_t = \varnothing$ and specify C and C*

Train SVM classifier with the working set $W^{(0)}$

Obtain the label vector of the unlabeled set V.

for i = 1 to T // T is the number of iterations

Select N+ positive transductive samples from the upper side of the margin and N− negative transductive samples from the lower side respectively.

Select positive candidate set B+ containing N+ positive transductive samples and negative candidate set B− containing N− negative transductive samples respectively.

$B^{(i)}{}_t = B^+ \cup B^-$

Update the training set:

if $A^{(i-1)}{}_t = \varnothing$ $W(i) = W^{(i-1)} \cup B^{(i)}{}_t$

$D^{(i)}{}_t = B^{(i)}{}_t$

Else

$D^{(i)}{}_t = A^{(i-1)} t \cap B(I) t$

$W^{(i)} = (W^{(i-1)} - D^{(i-1)}{}_t) \cup D^{(I)} t$

End if

$A^{(i)}{}_t = B^{(i)}{}_t$

Train TSVM classifier with the updated training set W (i)

Obtain the label vector of the unlabelled set V

end for

end

The designing of TSVM for the log file categorization is done in two ways and it addresses two issues. Those are

- Select the log file with the expected accurate labelling
- Choose the informative knowledge source about the log file

TSVM selects the margin of the classification of log file samples by accurately identifying the upper and lower side of the samples with analysing and checking the following conditions. Those are, if $P \geq 1$ (i.e., falls below or above the hyper plane region will be identified), then the transductive samples of log files closest to the margin bounds will be assigned as class labels as +1 or -1 respectively. Else labelling will be done any way without any consideration.

A dynamic adjustment is necessary, taking into account that the position of the hyper plane keeps changing in iterations. Typically, the most confident unlabelled patterns, together with their predicted labels, are added to the current training set. The classifier is retrained and the process is repeated. It is to be noted that the classifier uses its own prediction to teach itself. It is natural to imagine that a classification error can reinforce itself. Therefore, it is important to take a caution in the selection of transductive samples because wrong labelling may substantially degrade the performance of the classifier.

Due to the fact that support vectors contain the richest information among the informative samples (i.e., the ones in the margin band), the unlabelled patterns closest to the margin bounds have the highest probability to be correctly classified. Therefore, in this approach, design a selection procedure (i.e., filtering process) to increase the acceptability of the samples with the expected correct labelling. In other words, an unlabelled sample should be considered as transductive sample if the TSVM ensemble assigns the same label to it.

The above classification algorithm is used to result the categorization of the log files based on the behaviour of previous users. Thus with the knowledge of this algorithm one can categorize the log files based on the user's activity and behaviour. The attributes that are used for learning and categorizing whether the users are disengaged learners or not is listed in the Table 1.

Table 1: List of Attributes

S.No	Attributes
1	No of pages' read
2	Average time spent for Learning
3	Number of questions attended
4	Average time spent on assessment
5	Number of correct answers
6	Number of wrong answers

B. FIND THE DISENGAGED LEARNERS AMONG ACTIVE USERS

In the previous section, the learning of the previous online learners is done in terms of their disengagement activities. From the knowledge, the active users in the online learning system would be predicted and categorized as whether they are disengaged learner or active learner. This is done gathering the attributes values that are listed in the table 1 which would compare with the categorized labels of the training phase. The steps that are followed for categorization of actives users are listed as follows:

STEPS:

1. Gather the attribute values of the active users
2. Retrieve the categorized user information values
3. Compare the attribute values with the categorized label values
4. Predict disengaged learner or active learner

After analysing the active learners in the system as disengaged or active learners,

- The active learners would continue their process of online learning
- The disengaged learners would be classified into two classes namely partially disengaged or fully disengaged.

The disengaged learner classification process is explained clearly in the following sub sections.

C. CLASSIFY THE DISENGAGED LEARNERS BASED ON THE LEARNING ATTITUDE

The classification of disengaged learners is done to predict whether the learners are completely disengaged from the online learning system or they are distracted from the online learning system for the particular period of time. By knowing these, those learners can be assigned with tasks that most suit them. For example, the learners that are fully disengaged can be forwarded to attend an online survey and the partially disengaged learners can be handled by sending the popup message or doing small assessment tasks. In this work, naive bayes classification algorithm is used for classification of disengaged learners.

Naive Bayes is a simple technique for constructing classifiers: models that assign class labels to problem instances, represented as vectors of feature values, where the class labels are drawn from some finite set. It is not a single algorithm for training such classifiers, but a family of algorithms based on a common principle: all naive Bayes classifiers assume that the value of a particular feature is independent of the value of any other feature, given the class variable. For example, a fruit may be considered to be an apple if it is red, round, and about 10 cm in diameter. A naive Bayes classifier considers each of these features to contribute independently to the probability that this fruit is an apple, regardless of any possible correlations between the colour, roundness and diameter features.

For some types of probability models, naive Bayes classifiers can be trained very efficiently in a supervised learning setting. In many practical applications, parameter estimation for naive Bayes models uses the method of maximum likelihood; in other words, one can work with the naive Bayes model without accepting Bayesian probability or using any Bayesian methods.

```
Input trained data NB (C, D)  // C-class and D-Dataset
V← extract terms (D)
N← count dataset (D)
Find length of dataset (D/)
For each c
Compute probability of each sequence
Store probability in database.
Do count dataset in class (D, c)
Prior [c] ←
←  Concatenate from class distribution
```

Despite their naive design and apparently oversimplified assumptions, naive Bayes classifiers have worked quite well in many complex real-world situations. In 2004, an analysis of the Bayesian classification problem showed that there are sound theoretical reasons for the apparently implausible efficacy of naive Bayes classifiers. Still, a comprehensive comparison with other classification algorithms in 2006 showed that Bayes classification is outperformed by other approaches, such as boosted trees or random forests.

D. ASSIGN TASKS BASED ON THEIR CLASS

After classification of disengaged learners into fully disengaged and partially disengaged classes, some activities would be done to motivate the learners to continue learning through online learning system. The learners who does not have an interest on a specific course but joined due to the compulsion, improper material, poor network issues, improper communication with the instructor and the learner, whenever the learner needs an assistance with the system and the system fails to help the learner. Those reasons may lead the learner to fully disengaged on the online course. Hence the system will forward those disengaged learners to attend an online survey. The Online survey contains 10 set of questions. Based on the results from the online survey, an instructor has to identify the reason for fully disengaged and take necessary action to engage the learners. Table-2 describes the list of survey questions and possible answers for the questions. Similarly, identifying the reason for partially disengaged is also a tough case, because they are not completely disengaged. They are engaged most of the time and disengaged in a particular session. Whenever the learner seems to be demotivated, the system generate some additional assessment and send it to the learners to change their focus, create some pop up messages to change the user focus and so on. Similarly, if the learner is much engaged in learning activities but failed in examination activities, then the system forwards the learners to participate pre-tests before performing the actual examination. So that the learners will be flexible on assessment environment too.

Table 2: List of Survey Questions

Serial No	Survey Questions	Possible Answers
1	Are you Interest in the course?	{Yes, No}
2)	Whether the Online learning environment is attractive?	{Yes, No}
3)	Whether the online course material is well organized?	{Yes, No}
4)	Whether the language of the course material is understandable?	{Yes, No}
5)	Whether the online course material provides a sufficient example for each and every topic?	{Yes, No}
6)	The Online Course material belongs to which category?	{Basic, Intermediate, Advanced}
7)	How much time taken by the instructors to solve your queries?	{Immediate reply, Within a Day, Within a Week, Within a Month}
8)	Interaction among the other students	{Low, Not bad, High, More Positive}
9)	Does the use of Online learning system help you to improve your learning skills?	{Yes,No}
10)	How much u rate this online course?	{1- Very Poor, 2-Poor, 3- Average, 4- Good, 5- Very Informative)

IV. EXPERIMENTAL RESULTS

The experimental evaluation was conducted in the proposed research framework in which classification algorithm is used for discriminating the learning attitudes of the disengaged learners. In the proposed research, naive bayes algorithm is used for classification purpose which can provide improved results than the other classification algorithm. The evaluation is conducted between the various classification algorithms to prove the effectiveness of the proposed research methodology. The classification algorithms that are used in this work for experimental evaluation are listed as follows:

- Naive bayes
- Random Forest
- Random decision tree
- Alternative decision tree

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. Experimental results are listed in Table-3. Among the experimental results, we obtained, it confirms that the naïve bayes classification will give better classification results than other algorithms. The Confusion matrix of naïve bayes classification algorithm is listed in Table-4. The diagrammatic representation of the accuracy metric comparison is illustrated in the Figure 2.

Table 3: Experimental Results

	Naïve bayes	RF	RDT	ADT
% Correct	89.17	87.5	88.33	85
True Positive Rate	0.816	0.868	0.842	0.763
False Positive Rate	0.049	0.085	0.061	0.049
Precision	0.886	0.825	0.865	0.879
F-Measure	0.849	0.846	0.853	0.817
Error	0.109	0.125	0.117	0.15

Table 4: Confusion Matrix for Naive bayes Classification

	Predicted				
Actual		Engaged	Partially Disengaged	Fully Disengaged	Total
	Engaged	31	7	0	38
	Partially Disengaged	4	67	0	71
	Fully Disengaged	0	2	9	11
	Total	35	76	9	120

Figure 2. Accuracy comparison chart

In the Figure 2, accuracy comparison is done between the various classification algorithms namely Naive bayes, Random Forest, Random Tree and Alternative decision tree. From this figure, it is proved that the naive bayes algorithm can accurately discriminate the disengaged learner's behaviour then the other classification algorithms. It is because of; naive bayes algorithm consumes less number of iteration than the other classification algorithms. The accuracy value of the naive bayes algorithm is 89.17, Random Forest algorithm is 87.5, Random decision tree algorithm is 88.33 and Alternative decision tree is 85.

V. CONCLUSION

Online learning system is the most popular field in the real world environment such as teaching, learning, research field and so on. Learners are motivated to use the online learning system to avoid attending schools or libraries. In the proposed research work, the learner's attitude can be classified into three types namely Engaged, fully disengaged, partially disengaged. The most of the reason for fully disengaged lies outside of the scope of the log files, hence it is difficult to identify the solution. Thus we conducted an online survey that includes a questionnaire to know the user behavioural thoughts. Similarly, for the partially disengaged learners, whenever the learners lead to disengagement. The system generates pop up messages to change the user focus and so on. In the future perspective, we need to classify the partially disengaged learners in to more categories and motivate them based on their category.

VI. REFERENCE

1. 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.
2. Siragusa, L., & Dixon, K. (2008), Planned behaviour: Student attitudes towards the use of ICT interactions in higher education. In Hello! Where are you in the landscape of educational technology? Proceedings ascilite Melbourne 2008. Retrieved May 3, 2013
3. Gasaymeh, A. (2009), A study of faculty attitudes toward internet-based distance education: A survey of two Jordanian public universities. PhD Thesis, the College of Education, Ohio University, Ohio, USA.
4. Mishra, S., & Panda, S. (2007), Development and factor analysis of an instrument to measure faculty attitude towards e-learning. Asian Journal of Distance Education, 5(1), 27-33.
5. Wen, J. R., & Shih, W. L. (2008), Exploring the information literacy competence standards for elementary and high school teachers. Computers & Education, 50(3), 787-806.
6. 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.
7. Sundar PVP, Kumar AVS, Disengagement detection in online learning using Quasi framework. IOSR Journal of Engineering (IOSRJEN). 2015; 5(1):4–9.
8. 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.
9. 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.

10. Sundar PVP, Kumar AVS, A systematic approach to identify the unmotivated learners in online learning. Indian journal of science and technology. 2016. April; Vol 9(14): 1-6.

11. P. V. P. Sundar and A. V. S. Kumar, "A Novel disengagement detection strategy for online learning using quasi framework," Advance Computing Conference (IACC), 2015 IEEE International, Bangalore, 2015, pp. 634-638. doi: 10.1109/IADCC. 2015.7154784

12. Egbo, O. P., Okoyeuzu, C. R., Ifeanacho, I. C., & Onwumere, J. U. (2011), Gender perception and attitude towards e-learning: A case of business students, University of Nigeria. International Journal of Computer Application, 1(2), 135-148.

13. Nassoura, A. B. (2012),. Students' acceptance of mobile learning for higher education in Saudi Arabia. American Academic & Scholarly Research Journal, 4(2). Retrieved January 20, 2013

14. Bhuasiri, W., Xaymoungkhoun, O., Zo, H., Rho, J, J., & Ciganek, A, P. (2012), Critical success factors for e-learning in developing countries: A comparative analysis between ICT experts and faculty. Computers & Education, 58 (2012), 843-855

15. Deb, S. (2011), Effective long distance learning in developing countries using mobile and multimedia technology. International Journal of Multimedia and Ubiquitous Engineering, 6(2), 33-40.

16. Hussain, I. (2007), A study of student's attitude towards virtual education in Pakistan. Turkish Journal of Distance Learning, 8(2), 69-79. Retrieved February 26, 2013

17. Omidinia, S., Masrom, M., & Selamat, H. (2011), Review of e-learning and ICT infrastructure in developing countries (Case study of Iran). American Journal of Economics and Business Administration, 3(1), 120-125. Retrieved January 22, 2013

18. Selim, H. M. (2007), Critical success factors for e-learning acceptance: Confirmatory factor models. Computers and Education, 49(2007) 396–413. Retrieved May 3, 2013

19. Vrana, V., Fragidis, G., Zafiropoulos, C., & Paschaloudis, D. (nd), Analysing academic staff and students' attitudes towards the adoption of e-learning. Retrieved May 5, 2013.

20. Quasi Framework. Available from: www.quasiframework.com.