

Study of Behavioural Prediction in Online Courses

Bhagyashri Ganorkar¹, Anjali Raut²

¹Computer Science And Engineering, SGBAU,

²Computer Science And Engineering SGBAU,

Abstract— *The educational system is changing day by day. In that e-learning is trending this provides great platform and resources to students. Before online education, the focus is on long term courses where student performance cannot judge quickly. In E-learning the performance of the student is getting quickly. So it is easier to detect where the student is lacking by tracking the behaviour of them. Predictive learning analytics is used to improve the quality of online courses. Prediction is based on quiz scores, exam performance, and beneficial collaboration groups which help instructors to detect student performance which also helps students to increase their knowledge.*

Keywords— *Predictive Learning Analytics, Social Learning Network, Learning Outcomes Prediction, Behavioural Analysis, Performance Prediction.*

I. INTRODUCTION

In today's era, people attract things which give the fastest output at low price without any hard work. It may be shopping, entertainment or education. Where education is most important for the personal, social and economic development of the nation. As people always search for the best and quick result opportunities here the education also changes from the traditional classroom to the online to give the best and fast result. There are various online learning platforms are available with different teaching strategy. The main focus of e-learning system is on student assessment. For the assessment of student, the focus is on learner's behaviour. To analyse the behaviour of learner it focuses on student's drop-off rates, quiz scores, and exam performance and also how they interact with social learning network and course content. According to that system gives feedback to the students. As this action can take quickly the students can improve their performance fast. In feedback, students come to know their weak areas. There is another section of online learning is the discussion forum. In a discussion forum student can discuss their problems and anyone can reply on it. To improve the quality of course predictive learning analytics is used.

II. LITERATURE REVIEW

This work is related to data mining and machine learning for education. Here the discussion is on learning outcomes prediction.

a. Learning Outcomes Prediction:

W. H'am'al'ainen and M. Vinni published a paper "Classifiers for educational data mining" which is focused on the prediction of student performance. The predictive learning analytics is used to predict the behaviour of learner in advance mainly students in the online course [1]. It track the learner's activity such as performance of learners on assessments [2],[3], [4], [5] learners risk of obtaining unfavourable outcomes [6], [7] learners final grades [8], [9], [10], [11] and learners hard work[12], [13], [9].

Another perspective is to focus on predicting binary outcomes in short courses which are characterized by timescales and a lack of intermediate quiz/assessment data. If an assessment is not done properly then quiz-based outcomes prediction is not applicable directly. Instead, our method turns learners' content clickstreams and social learning data into features for binary outcome classification.

b. Content Feature Mining:

Student's clickstream data can be considered for behavioural mining by the researcher. Researchers have proposed some methods which focus on the analysis of clickstream which count the number of clicks by students every day. Grade prediction in online courses focuses on video watching behaviour [14], [12], [2] which defines aggregate quantities like a fraction of time spent and the number of rewinds.

[13] Clickstream feature is used for outcome prediction such as mining learner's activity, forum, wikis, and quiz prediction and [6] employed mixture models to group students based on time spent for predicting certification. This work is [2], [6] related to clickstream data for early detection. The short time period for courses may be affected to additional modelling challenges that we overcome by defining features for each piece of content separately.

Day by day prediction helps the instructor to predict performance on the basis of learners revisited contents rather than predicting unit by unit scheme [2]. [6] In this authors also study for early detection of binary outcomes, with defining many features beyond time spent and the number of rewinds to find more prediction in short-courses.

c. Social Learning Feature:

The several studies are considered the social learning networks in different learning platform such as MOOC[15], [9], [16], online courses[11], [17], [18], Q and A sites[19], and enterprise social networks[20]. Similar to the discussion forum in online courses [15], users asking and answering questions emerged the social learning network on Q and A sites, so the analysis method is overlap in these scenarios.

The prior work is to focus on SLN itself; such as exploratory analysis [19], [20] and optimization of interactions [15]. According to this [9] graphical model is developed to predict grade and completion from learner post and reply frequencies,[11] and also make the prediction of final performance from participation indicates in both quantitative, qualitative and social network forums. In addition to social learning network, our method defined other types of features which are not present in prior work. The additional feature contains topic similarity between learner's posts and course content feature that how they interact with course material. In our short-courses, we find some observations which are topic similarity features are particularly predictive and content feature which is useful for early detection.

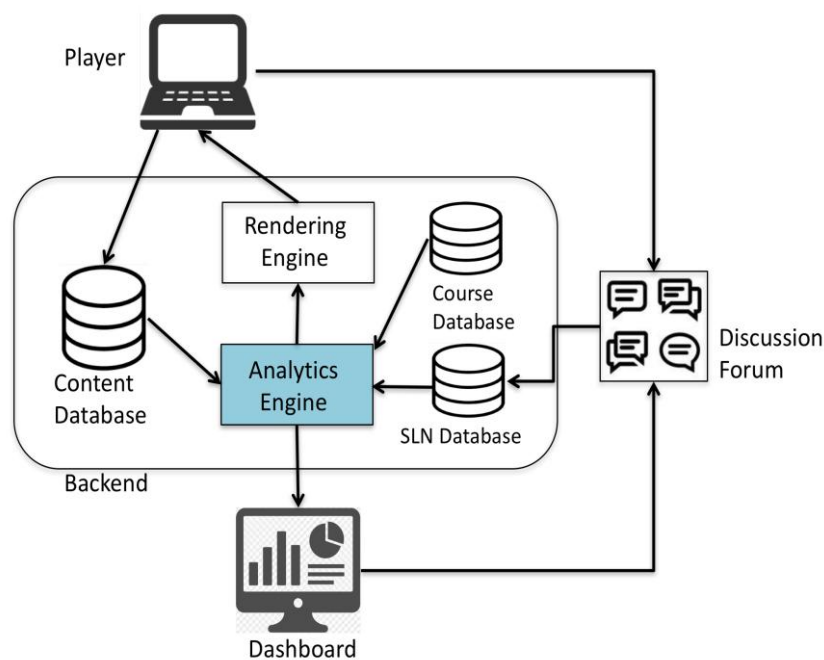


Fig. 1 System Architecture Overview

A. Predictive Learning Analytics:

Predictive learning analytics is emerging as a research area and it is used to improve course quality mainly in online courses. Generally, predictive learning analytics is developed for semester-long courses for example- Massively Open Online Courses. This course scenario found two properties which is useful for modelling perspective. First one is frequent assessments to track the student progress and the second one is a large number of enrolled learners. But what happens when we consider the following cases: 1) when assessment of student does not use frequently and 2) when there is a small number of the learner's enrolled. Therefore these short online courses are required to develop the predictive learning analytics methodology in such a way that it works in the entire situation.

In today's online courses generally, it tracks the behaviour of learner such as their interaction with social learning network and with the course content. According to that design a PLA method that model learner attribute based on behavioural data in online short courses.

B. Behaviour-Based Outcomes Prediction:

Here we find out some questions related to learning outcome prediction

- 1) *Can behavior alone be used to predict the learning outcomes?*
- 2) *How quickly this prediction made in an online course with the quality?*
- 3) *Learning behavior associated with SLN and course content can be a better way for prediction?*

An algorithm is developed by researchers for student performance prediction. As we consider the challenge that comes in short courses for example small enrolment and low assessment for that one methodology are developed. To make the prediction of nth days first we have to collect the behavioural data that is from the first n day's prior offerings of this course as an input. Here the challenge is to process raw data into effective features sets for modelling learning behaviour. *There are two types of features are as follows:*

- 1) *Course Content:* It includes learner behaviour while interacting with course content
- 2) *SLN features:* It concludes learner's discussion in the forum. It includes a similarity between learners post and course content.

The objective is to predict that the learners will pass or fail the course our method uses the feature set as input to different classifiers in training and evaluation.

III. PROPOSED WORK

In the proposed project, the focus is on short term online courses and personalized study material recommendation for students depending on their understanding level. In this system an algorithm is proposed which is a deep learning algorithm to detect weak subject/topic of students by using the learning behaviour of the student. The learning behaviour will be tracked while using the forum, appearing topic wise tests, etc. another concept which is introduced is online group study to improve student's performance. The system will automatically detect weak and strong topics of student and depending on this; our user-defined algorithm will suggest group members to students accordingly. The group circle must contain some strong and some weak students so that strong students will solve weak student's doubts. After weak topics detection, the system will automatically recommend study material for weak topics and recommend members for group studies. Students have to improve their weak topic with the help of allotted study material and group discussion.

The system contains the following modules:

- 1) *Admin panel:* In this module, admin add the tutors, manage the course and view the student performance.
- 2) *Tutor panel:* In this module, the tutor can participate in the forum, can view students' performance.
- 3) *Student:* In this module, the student can subscribe for a course, view the study material, solve the test, view their own performance, take a part in the forum, can see their group members.
- 4) *Study group formation:* In this module, students group will be formed according to their ability.
- 5) *Weak Subject/topic Detection:* In this module, system track student's activity and according to that find out the weak topic of students.
- 6) *Study material recommendation:* According to the student's IQ and weak subjects, the system will study material automatically to the student.

IV. CONCLUSION

The technique is used to predict learners' performance from their behaviour in online short courses. It is challenging to make the prediction about learners with lack of intermediate assessment in small enrolment. The method depends on machine learning behaviour which is obtained by measurements processing during the learning process and learners' interaction with content in learning networks. When collected data is processed under the different conditions then high prediction quality in the middle, of course, will be achieved, focusing on the methods to provide early detection of performance. We found that the social learning network is more helpful for prediction but require large time where the content attribute gives the better quality for early detection in a short time. Next, we found that the method can give behavioural analysis to the instructor.

REFERENCES

- [1] W. H'am'al'ainen and M. Vinni, "Classifiers for educational data mining," Handbook of Educational Data Mining, pp. 57–74, 2010.
- [2] C. G. Brinton and M. Chiang, "MOOC Performance Prediction via Clickstream Data and Social Learning Networks," in IEEE INFOCOM, 2015, pp. 2299–2307.
- [3] T.-Y. Yang, C. G. Brinton, C. Joe-Wong, and M. Chiang, "Behavior based grade prediction for moocs via time series neural networks," IEEE JSTSP, 2017.
- [4] A. S. Lan, C. Studer, and R. G. Baraniuk, "Time-varying learning and content analytics via sparse factor analysis," in SIGKDD. ACM, 2014, pp. 452–461.

- [5] *C. Piech, J. Bassen, J. Huang, S. Ganguli, M. Sahami, L. J. Guibas, and J. Sohl-Dickstein*, “Deep Knowledge Tracing,” in NIPS, 2015, pp. 505–513.
- [6] *C. A. Coleman, D. T. Seaton, and I. Chuang*, “Probabilistic use cases: Discovering behavioral patterns for predicting certification,” in L@S. ACM, 2015, pp. 141–148.
- [7] *H. Lakkaraju, E. Aguiar, C. Shan, D. Miller, N. Bhanpuri, R. Ghani, and K. Addison*, “A Machine Learning Framework to Identify Students at Risk of Adverse Academic Outcomes,” in ACM SIGKDD, 2015, pp. 1909–1918.
- [8] *Y. Meier, J. Xu, O. Atan, and M. van der Schaar*, “Predicting Grades,” IEEE Trans. Signal Proc., vol. 64, no. 4, pp. 959–972, 2016.
- [9] *J. Qiu, J. Tang, T. X. Liu, J. Gong, C. Zhang, Q. Zhang, and Y. Xue*, “Modeling and Predicting Learning Behavior in MOOCs,” in ACM WSDM, 2016, pp. 93–102.
- [10] *B. Minaei-Bidgoli, D. A. Kashy, G. Kortemeyer, and W. Punch*, “Predicting student performance: an application of data mining methods with an educational web-based system,” in Frontiers in education, 2003., vol. 1. IEEE, 2003, pp. T2A–13.
- [11] *C. Romero, M. I. L’opez, J. M. Luna, and S. Ventura*, “Predicting Students’ Final Performance from Participation in Online Discussion Forums,” Computers & Education, vol. 68, pp. 458–472, 2013.
- [12] *T. Sinha, P. Jermann, N. Li, and P. Dillenbourg*, “Your Click Decides Your Fate,” in ACL EMNLP, 2014, pp. 3–14.
- [13] *J.-Y. Lee, A. Toescher, M. Jahrer, K. Ozaki, M. Bay, P. Yan, S. Chen, T. T. Nguyen, and X. Zhou*, “Multi-stage ensemble and feature engineering for mooc dropout prediction,” June 2016. [Online]. Available: <http://www.conversionlogic.com/>
- [14] *C. G. Brinton, S. Buccapatnam, M. Chiang, and H. V. Poor*, “Mining MOOC Clickstreams: Video-Watching Behavior vs. In-Video Quiz Performance,” IEEE TSP., vol. 64, no. 14, pp. 3677–3692, 2016.
- [15] *C. G. Brinton, S. Buccapatnam, F. M. F. Wong, M. Chiang, and H. V. Poor*, “Social Learning Networks: Efficiency Optimization for MOOC Forums,” in IEEE INFOCOM, 2016.
- [16] *M. Kumar, M.-Y. Kan, B. C. Tan, and K. Ragupathi*, “Learning Instructor Intervention from MOOC Forums: Early Results and Issues,” ERIC EDM, pp. 218–225, 2015.
- [17] *H.-L. Yang and J.-H. Tang*, “Effects of social network on students’ performance: a web-based forum study in taiwan,” Journal of Asynchronous Learning Networks, vol. 7, no. 3, pp. 93–107, 2003.
- [18] *M. Abdous, H. Wu, and C.-J. Yen*, “Using data mining for predicting relationships between online question theme and final grade,” Journal of Educational Technology & Society, vol. 15, no. 3, p. 77, 2012.
- [19] *C. G. Brinton and M. Chiang*, “Social Learning Networks: A Brief Survey,” in IEEE CISS, 2014, pp. 1–6.
- [20] *J. Cao, H. Gao, L. E. Li, and B. Friedman*, “Enterprise Social Network Analysis and Modeling: A Tale of Two Graphs,” in IEEE INFOCOM, 2013, pp. 2382–2390.