International Journal of Scientific Engineering and Research (IJSER)
ISSN (Online): 2347-3878
Index Copernicus Value (2015): 56.67 | Impact Factor (2017): 5.156

Student Performance Prediction for Education

Radhika Kale

Department of Computer Engineering, VPKBIET, Baramati, India
radhikakale145 [at]gmail.com

Abstract: ***To assure students' graduation on time and that too with satisfactory performance, the accurate prediction of future performance based on their ongoing academic details is needed. Very few searches have been made for predicting students' performance in completion of degrees as most searches focuses performance prediction while solving problems. The paper introduces a new way which uses machine learning for prediction of student's performance in degree program which is able to counter important challenges that other systems faces like difference in students' backgrounds and selected courses, inequality in an information that courses provide for predictions etc.***

Keywords: Prediction of Student performance, data-driven course clustering, personalized education

1. Introduction

Data mining has made its applications spread all across educational domain and is being very helpful. For better intervention, it is necessary to build a system that can keep track of students' performance and predict students' future performance accurately. Existing approaches shows difficulties in covering diversities of educational backgrounds that student's come from and students' current performance is not taken into an account as well. Also, how much information a set of course can provide is not considered here as well. To counter these drawbacks this paper proposes a system that covers all these drawbacks. System consists of two layers of predictors namely base predictors and ensemble predictors for performance predictions. This system also clusterizes the courses for finding relevant courses for training data according to student and the information that a student provides.

2. Literature Survey

Rahel Bekele and Wolfgang Menzel proposed the importance of accurate estimation of student's future performance is essential in order to provide the student with adequate assistance in the learning process. This research aimed at investigating the use of Bayesian networks for predicting performance of a student, based on values of some identified attributes [2].C. MARQUEZ-VERA proposed to apply data mining techniques to predict school failure. Several experiments have been carried out in an attempt to improve accuracy in the prediction of final student performance and specially of which students might fail [3]. H. Cen, K. Koedinger, and B. Junker proposed a semi-automated method for improving a cognitive model called Learning Factors Analysis that combines a statistical model, human expertise and a combinatorial search. This method is used to evaluate an existing cognitive model and to generate and evaluate alternative models. To use the method for datasets from other tutors to discover its potential for model and tutor improvement [4]. Nguyen Thai-Nghe, Tomas Horvath and Lars Schmidt- Thieme proposes to take into account the sequential effect, this work proposes a novel approach which uses tensor factorization for forecasting student performance [5]. M.Feng, N.Heffernan and Kenneth Koedinger proposed the assessing student math proficiency is to use data that our system collects through its interactions with students to estimate their performance on an end-of-year high stakes state test. This result show that we can do a reliably better job predicting student end-of-year exam scores by leveraging the interaction data, and the model based on only the interaction [6]. Man-Ching Yuen, Irwin King, Kwong Sak Leung proposed a Task Recommendation (TaskRec) framework based on a united probabilistic matrix factorization, aiming to recommend tasks to workers in dynamic scenarios. Unlike traditional recommendation systems, workers do not provide their ratings on tasks in crowd sourcing systems, thus they infer user ratings from their interacting behaviors [7].

3. Methodology

1. Ensemble-based Progressive Prediction (EPP):

This layer does the noise reduction part and finds relevant courses for student's performance prediction. The predicted output is then again passed to the base predictors for adaptation and consideration of student's performance state in future prediction. The ensemble predictor synthesizes the previous ensemble output and output of the base predictors and makes a final prediction. The ensemble predictor is learned using student data. Learning the ensemble predictors is done online. The output predicted by ensemble layer is further passed to the base predictor layer. This ensures that the current performance state of student is taken into consideration for next prediction.

2. Data Driven approach:

A data driven approach is used to utilizing the educational domain knowledge. This approach is used to learn the relevant courses. This method is based on probabilistic matrix factorization.

4. Conclusion

A new unique technique for predicting students' performance in degree programs. This is done by using their current and previous performances. A latent factorization model based clustering method for courses is built for retrieving relevant courses. Only these relevant course's records are provided as a training data to prediction model. This reduces the

complexity of training as well as irrelevant data. This leads in better performance in the prediction. These methods are or can be used along with pedogeological methods to take necessary actions in order to make students' performance better.

References

[1] Jie Xu, Member, IEEE, Kyeong Ho Moon and Mihaela van der Schaar "A Machine Learning Approach for Tracking and Predicting Student Performance in Degree Programs"DOI10.1109/JSTSP.2017.2692560IEEE Journal.
[2] R. Bekele and W. Menzel, "A Bayesian approach to predict performance of a student (bapps): A case with Ethiopian students", algorithms, vol. 22, no. 23, p. 24, 2005.
[3] C. Marquez-Vera, C. Romero, and S. Ventura, Predicting school failure using data mining, in Educational Data Mining 2011, 2010.
[4] H. Cen, K. Koedinger, and B. Junker, Learning factors analysis a general method for cognitive model evaluation and improvement, in Inter National Conference on Intelligent Tutoring Systems. Springer, 2006, pp.164175.
[5] N. Thai-Nghe, T. Horvath, and L. Schmidt-Thieme, Factorization models for forecasting student performance, in Educational Data Mining2011, 2010.
[6] M. Feng, N. Heernan, and K. Koedinger, Addressing the assessment Challenge with an online system that tutors as it assesses, User Modeling And User-Adapted Interaction, vol. 19, no. 3, pp. 243266, 2009.
[7] M.-C. Yuen, I. King, and K.-S. Leung, "Task recommendation in crowd-sourcing systems", in Proceedings of the First International Workshop o Crowd sourcing and Data Mining. ACM, 2012, pp. 2226.

Author Profile

Radhika Kale pursuing the M.E degrees in Computer Engineering from Vidya Pratishthan Kamalnayan Bajaj Institute of Engineering and Technology.