A Machine Learning Approach for Tracking and Predicting Student Performance in Degree Programs

Deepthi #1, Mr J Rakesh #2
#1 M.Tech., Scholar, #2 Asst. Professor,
Computer Science and Engineering,
Priyadarshini Institute of Technology and Management

ABSTRACT

Predicting students' future performance based on their current academic records is critical to ensuring that students graduate on time and to their satisfaction. Even though there is a lot of research on how to forecast student achievement in solving issues or preparing for courses, there is little research on how to predict student performance in finishing degrees (e.g. college programmes). Students come from a wide range of backgrounds and take a variety of courses, so it's impossible to predict how well they'll perform based just on what they've studied. New machine learning methods for forecasting student success in degree programmes have been developed in this study to solve these fundamental issues. Among the method's main characteristics are two. In order to make predictions based on students' changing performance states, a bilayered structure with numerous base predictors and a cascade of ensemble predictors is created first. We also suggest using latent component models and probabilistic matrix factorization to find out if a course has any importance in terms of building effective base predictors. We demonstrate that the proposed strategy outperforms benchmark methods via extensive simulations using an undergraduate student dataset gathered over three years at UCLA

1. INTRODUCTION

The government places a great priority on supporting the economic well-being of the country by making higher education more accessible. There is still a trillion-dollar student loan debt in America, which is more than the combined total of Americans' credit card and vehicle loan bills. The ever-increasing cost of college education (tuitions, fees, and living costs) has made it more difficult for students to graduate in a timely manner. Only 50 of the more than 580 public four-year universities in the United States have on-time graduation rates exceeding 50% for full-time students. To make college more affordable, it is necessary to guarantee that many more students graduate on time by early interventions on students whose performance will not match the graduation standards of the degree programme on time. Systematic tracking of student development and precise prediction of future outcomes, such as when they are expected to graduate and how high their GPAs may be, is a vital step toward successful intervention. When it comes to studying student performance, there has been a lot of research done, but it has been mostly focused on ITSs and MOOCs, where students solve issues in a classroom or on a platform like Massive Open Online Courses (MOOCs).

There are considerable differences and obstacles when it comes to forecasting the success of students in a degree programme (such as a college programme). First, students' backgrounds and areas of interest (majors, specialties) might vary greatly, resulting in a wide range of course choices and sequences; on the other hand, the same course can be taken by students from various backgrounds. In order to train a successful predictor, one must be able to deal with diverse student data owing to varied areas and interests, which makes it difficult to forecast student performance in a certain course. Students in ITSs, on the other hand, are more likely to follow a set of standard procedures while addressing issues. In the same way, in-class evaluations are sometimes used to forecast students' achievement in a course. There are various courses that students may take, but not all of these courses are equally useful in forecasting a student's future academic success. In addition to increasing the forecast's complexity, including the student's prior performance in all of his or her courses results in a worse prediction since it adds noise into the model. For example, a student's grade in "Linear Algebra" can be used to forecast a student's grade in "Linear Optimization," whereas a student's mark in "Chemistry Lab" may have considerably less predictive ability. It's not always as clear-cut as in this situation, however. If you want precise estimates of student performance, you must find out what courses are related. As students go through their degree programme, it is not possible to estimate their future performance on a one-time basis. Instead, it is necessary to keep track of how they are doing and to update their predictions as they complete new courses. Accordingly, the forecast has to be based on more than just the most current snapshot of student successes, but also on how progress has evolved over time, which may provide vital information for creating more accurate predictions. However, the complexity may readily grow since simply calculating the progression of student growth can be a difficult operation.

When forecasting the future, it may not be a good idea to regard the past as equally important as the present performance, because intuition tells us that old knowledge likely to be out of date. Here, we provide an innovative strategy for forecasting student success in degree programmes, taking into account the limitations outlined above. Although our primary emphasis is on forecasting GPAs, the fundamental architecture may be applied to a variety of different tasks involving student performance. Our primary contributions are three-fold.

Predictions are made using a new algorithm we devised based on students' changing performance levels. An ensemble predictor layer and a base predictor layer form the bilayered structure. Each academic term, a set of basic predictors create local forecasts based on a snapshot of the student's present performance level. During the ensemble layer, an ensemble predictor synthesises both local forecasts and the previous-term ensemble prediction to provide a prediction of future performance. It is possible to include students' progress into the forecast while keeping the complexity modest by cascading the ensemble predictor across academic terms. Our suggested algorithm also has a performance guarantee.

Probabilistic matrix factorization (PMF) may be used to automatically cluster courses based on huge, diverse, and sparse student course grade information. Based on the course clustering findings, machine learning methods are used to train base predictors. For the base predictors, only relevant courses from the same cluster are input. Training complexity is reduced while also removing unnecessary information and reducing noise in prediction. The UCLA Mechanical and Aerospace Engineering Department's dataset of 1169 undergraduate students was gathered over three years and subjected to rigorous simulation. It is clear from the findings that our suggested strategy outperforms the benchmark methods while maintaining educational interpretability.

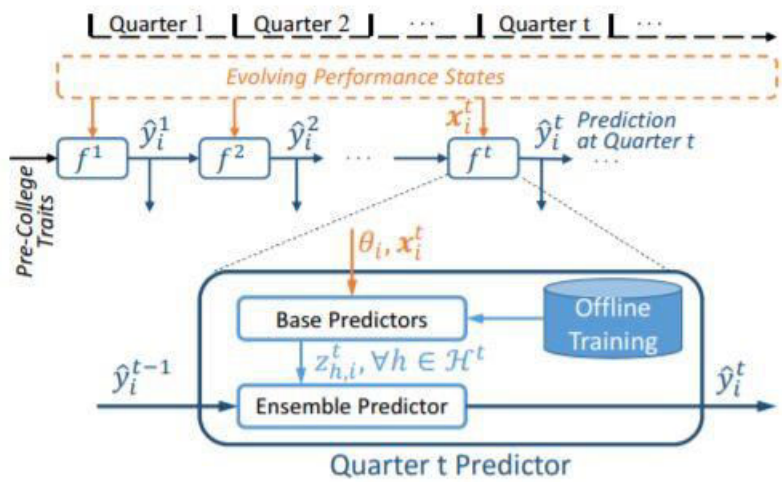

Fig.1 System Architecture

2. LITERATURE REVIEW

Learning factors analysis–a general method for cognitive model evaluation and improvement

An intelligent tutor's cognitive model is a collection of production rules or abilities that represent how pupils approach problem solving. Subject specialists, cognitive scientists, and programmers collaborate in brainstorming and iterative refinement to produce it. To improve our cognitive model called Learning Factors Analysis, we have developed an automated technique that combines a statistical model, human knowledge, and a combinatorial search. A cognitive model may be evaluated using this approach and alternative models can be generated and evaluated using this method. We offer new cognitive models and recommend improvements to the intelligent tutor based on them.

Addressing the assessment challenge with an online system that tutors as it assesses

Increasingly, secondary instructors in the United States are being requested to utilise formative assessment data (Black & Wiliam, 1998a, 1998b; Roediger & Karpicke, 2006) to guide classroom teaching. It is also being called "No Child Left Untested" by opponents of No Child Left Behind law in the United States. Critics point out that for every hour spent evaluating kids, an hour of education is missed. Then then, does it really have to be this way?

What if we enabled students to learn while taking the exam and better incorporated assessment into classroom instruction? Students who are unable to answer practise assessment problems on their own may take advantage of our new method, which gives them with quick help. Instead of only looking at whether or not students answer test questions correctly or incorrect, we believe that data on the amount of effort it takes for students to solve a test question with instruction might help us obtain more accurate results. The ASSISTment system incorporates both support and evaluation. By giving education to students and providing instructors with a more complete assessment of students' skills, the system allows teachers to make better use of their time. The data that our system obtains via its contacts with kids is used to predict students' performance on a high stakes state exam at the end of the year. Students' end-of-year exam scores are better predicted by using interaction data, according to our findings, and the model built only around this data outperforms the usual assessment model, which focuses on test question accuracy alone.

Personalized grade prediction: A data mining approach

Automated technologies assisting the lecturer are essential for both regular classroom courses and Massive Open Online Courses. One of the most difficult challenges is identifying students who are at risk of failing a class early enough to implement corrective measures. Each student in a class's final grade may be predicted using this approach. When the estimated accuracy of the forecast is sufficient, it provides a prediction for each student separately. Students' prior performance in a course helps the algorithm figure out the best time and place to make a prediction. Our approach is tested on a dataset of 700 undergraduate students who have completed an introductory digital signal processing course in the last seven years. Based on the results of a pilot course, our research implies that early in-class evaluations, such as quizzes, may be used to accurately forecast each student's performance, allowing the teacher to intervene as needed. Index Terms—Algorithms for forecasting, online education, prediction of grades, data mining, education in digital signal processing.

Mooc performance prediction via clickstream data and social learning networks

Predicting a student's performance in Massive Open Online Courses (MOOCs) is a focus of our research, which aims to determine if a user will answer a question Correct on First Attempt (CFA). As a result, we build new methods for analysing MOOC platform behavioural data. First, we extract summary quantities (such as the fraction played and number of pauses) for each user-video pair in our MOOC and show how certain intervals or sets of values for these behaviours quantify whether or not a pair is more likely to be CFA for the associated question.. or not.. for that question. Our approaches are based on these results and are meant to utilise training data intervals and the related success estimations as learning features in prediction algorithms to choose the best intervals for training. For all datasets and metrics studied, our methods exceed ordinary algorithms (i.e., without behavioural data) in terms of their performance. This "early detection" capabilities of clickstream data is especially evident when looking at the first few weeks of the course. CFA prediction may also be used to create graphs of students' Social Learning Networks (SLNs), which can assist teachers better manage their courses.

Data mining for adaptive learning in a tesl-based e-learning system

For an e-learning system, this research presents an adaptive learning in TESL (adaptive learning in teaching English as a second language), which incorporates different student characteristics. The AL-TESL-e-learning system uses an artificial neural network (ANN) to analyse the learning performance of diverse pupils. There are three tiers of material in the AL-TESL e-learning system that cater to various learning styles: vocabulary, grammar, and reading. An investigation comparing the AL-TESL-e-learning system outcomes to those of a standard online course control group sheds light on whether the suggested system can really be implemented. The AL-TESL-e-learning system outperformed a standard online course in terms of student learning performance, according to the findings of the statistical analysis.

3. SYSTEM ANALYSIS

Existing System

In reality, recent surveys reveal that just 50 of the more than 580 public four-year schools in the United States have on-time graduation rates at or over 50 percent for their full-time students . To make college more affordable, it is therefore necessary to guarantee that many more students graduate on time by early interventions on students whose performance would be unlikely to achieve the graduation standards of the degree programme on time. A vital step towards successful intervention is to design a system that can continually keep track of students' academic achievement and reliably anticipate their future performance, such as when they are likely to graduate and their predicted final GPAs, given the present progress. Although forecasting student performance has been widely examined in the literature, it was mostly addressed in the settings of answering problems in Intelligent Tutoring Systems (ITSs) (ITSs).

Detriments

On the other hand, forecasting student success in a degree programme (e.g., college programme) is fundamentally different and confronts new obstacles.

Proposed system:

A degree programme is one in which students must finish a series of courses in order to graduate after T semesters of formal education has passed. In order to take a course, students must first complete and pass a series of preparatory courses. A directed acyclic graph (DAG) may be used to represent the dependence between the prerequisites. Students may be required to finish a number of distinct subsets of courses in order to graduate from a particular programme. This department's prediction challenge will be the centre of our attention. However, we will continue to use data from other sources to help us make predictions. Data from a single location is typically restricted, although many courses are shared by many regions.

Advantages of proposed system

It is necessary for the development of effective foundational predictors. A system that is able to constantly monitor and properly anticipate the academic progress of kids.

4. ALGORITHMS

k-Means Clustering: k-means is one of the simplest unsupervised learning algorithms that solve the well known clustering problem. The procedure follows a simple and easy way to classify a given dataset through a certain number of clusters (assume k clusters) fixed apriori. The main idea is to define k centers, one for each cluster. These centers should be placed in a cunning way because of different location causes different result. So, the scheme, and that the sparse configuration and rank one significantly improves the performance of the recommendation. Better choice is to place the as much as possible far away from each other. The next step is to take each point belonging to a given data set and associate it to the nearest center. When no point is pending, the first step is completed and an early group age is done. At this point we need to re-calculate k new centroids as barycenter of the clusters resulting from the previous step. After we have these k new centroids, a new binding has to be done between the same data set points and the nearest new center. A loop has been generated. As a result of this loop we may notice that the k centers change their location step by step until no more changes are done or in other words centers do not move any more. Finally, this algorithm aims at minimizing an objective function know as squared error function given by

$$J(V) = \sum_{i=1}^{c} \sum_{j=1}^{c_i} (\|x_i - v_j\|)^2$$

SVM: In data analytics or decision sciences most of the time we come across the situations where we need to classify our data based on a certain dependent variable. To support the solution for this need there are multiple techniques which can be applied; Logistic Regression, Random Forest Algorithm, Bayesian Algorithm are a few to name. SVM is a machine learning technique to separate data which tries to maximize the gap between the categories.

5. RESULTS

Fig 2: Application home page

Fig 3: Uploading datasat

Fig 4: Finding the mean square error (MSE)

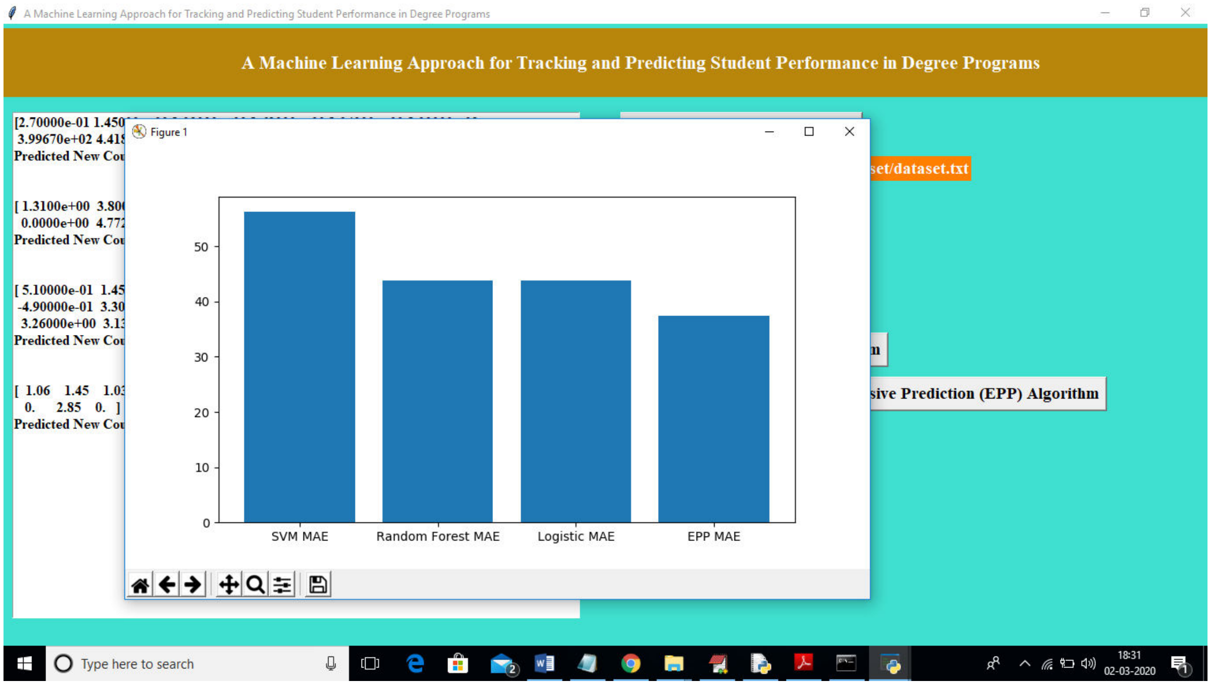

Fig 5: MSE Graph using different algorithms

6. CONCLUSION

Students' future success in degree programmes may be predicted using their present and previous performance, according to a new approach suggested in this research. To find relevant courses for developing base predictors, researchers devised a course clustering technique based on a latent component model. As students' performance progresses, a new ensemble-based progressive prediction architecture was designed. With the use of data-driven methodologies, academic advisers may propose future courses and carry out pedagogical intervention measures if appropriate, based on the results of the evaluations of students' performance. Degree programme curricula and education policy are also likely to be influenced by this research. Predicting student success in optional courses and utilising that information to help students choose course selections will be the focus of future research.

REFERENCES

[1] The White House, “Making college affordable,” https:// www.whitehouse.gov/issues/education/higher-education/ making-college-affordable, 2016.

[2] Complete College America, “Four-year myth: Making college more affordable,” http://completecollege.org/wp-content/uploads/2014/ 11/4-Year-Myth.pdf, 2014.

[3] H. Cen, K. Koedinger, and B. Junker, “Learning factors analysis–a general method for cognitive model evaluation and improvement,” in International Conference on Intelligent Tutoring Systems. Springer, 2006, pp. 164–175.

[4] M. Feng, N. Heffernan, 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. 243–266, 2009.

[5] H.-F. Yu, H.-Y. Lo, H.-P. Hsieh, J.-K. Lou, T. G. McKenzie, J.-W. Chou, P.-H. Chung, C.-H. Ho, C.-F. Chang, Y.-H. Wei et al., “Feature engineering and classifier ensemble for kdd cup 2010,” in Proceedings of the KDD Cup 2010 Workshop, 2010, pp. 1–16.

[6] Z. A. Pardos and N. T. Heffernan, “Using hmms and bagged decision trees to leverage rich features of user and skill from an intelligent tutoring system dataset,” Journal of Machine Learning Research W & CP, 2010.

[7] Y. Meier, J. Xu, O. Atan, and M. van der Schaar, “Personalized grade prediction: A data mining approach,” in Data Mining (ICDM), 2015 IEEE International Conference on. IEEE, 2015, pp. 907–912.

[8] C. G. Brinton and M. Chiang, “Mooc performance prediction via clickstream data and social learning networks,” in 2015 IEEE Conference on Computer Communications (INFOCOM). IEEE, 2015, pp. 2299– 2307.

[9] KDD Cup, “Educational data minding challenge,” https://pslcdatashop. web.cmu.edu/KDDCup/, 2010.

[10] Y. Jiang, R. S. Baker, L. Paquette, M. San Pedro, and N. T. Heffernan, “Learning, moment-by-moment and over the long term,” in International Conference on Artificial Intelligence in Education. Springer, 2015, pp. 654–657.

[11] C. Marquez-Vera, C. Romero, and S. Ventura, “Predicting school failure using data mining,” in Educational Data Mining 2011, 2010.

[12] Y.-h. Wang and H.-C. Liao, “Data mining for adaptive learning in a tesl-based e-learning system,” Expert Systems with Applications, vol. 38, no. 6, pp. 6480–6485, 2011.

[13] N. Thai-Nghe, L. Drumond, T. Horvath, L. Schmidt-Thieme ´ et al., “Multi-relational factorization models for predicting student performance,” in Proc. of the KDD Workshop on Knowledge Discovery in Educational Data. Citeseer, 2011.

[14] A. Toscher and M. Jahrer, “Collaborative filtering applied to educational data mining,” KDD cup, 2010.

[15] 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.