STUDENTS ON TRACK

Early Warning System

Danny Green and Le Van Anh
RMIT University Vietnam

Rationale

This project was initialised for two main reasons. The first is that the current at-risk classification system at the university is based on the amount of courses a student fails compared to their course load. Many, including those working on this project, view this as identifying poor preforming students after the fact, making interventions and support services ineffective. The second is that there is little institutional research investigating achievement indicators in the unique context of the university, which is one where almost the entire student body is studying in a foreign language (English) despite being in their home country and sharing the same first language. Additionally, as this is a private university, the only students that could be traditionally classified as being of low socioeconomic status are the very few who attend through scholarships. Thus there was a need to build a timelier system based on a more context appropriate model

Existing At-Risk System

The existing At-Risk system is based on courses failed compared to the students course load. If a student fails 50% of their course load, they are deemed officially At-Risk. To illustrate, if a student is studying 2 courses and fails 1, the become At-Risk 1. This is flagged in the Student Record System (SRS). Furthermore, if this occurs again, they then become At-Risk 2, and likewise for At-Risk 3. At-Risk 2 and 3, they are case managed by the university's student advisement team, who offer one to one advisement and performance improvement plans and refer students to appropriate support such as academic advisement or health and wellbeing when necessary. There is a major drawback to this system, which is that cases are often only referred to academic support in mid to late semester, making it difficult for this service to be effective.

The General Model:

Looking at over 78 000 student records across 2 years with over 20 variables, the researchers built an initial model using multiple linear regression and Bayesian Model Averaging (BMA) through the statistical computing language R to try to predict student achievement . The model identified factors such as high school grade averages, English level, and the amount of credit points achieved in the program and previous academic performance as the most significant factors influencing achievement. This model had an R-squared of 0.6815 and in testing, 88% of the students who failed were predicted by the model's "focus-list".

The Early Warning System

System Pilot, Semester 3 , 2016

The pilot model:

The pilot involved three first-year, high enrolment, core business courses: Introduction to Management , Marketing Principles and Business Computing, which were selected by the University's At-Risk Committee. Researchers then refined the model to be specifically relevant to the chosen courses by selecting a sample formed of records only from those three courses. A training dataset was then created (ratio = 7:3) and BMA was applied to ascertain the most significant predictors of final grades: average high school score, English proficiency, credit points achieved, current GPA, whether they were repeating the course, the number of student advisement notes and their at-risk status. Multiple Linear regression was then applied using these predictors to develop a predictive model.

The model has a multiple R-squared of 0.416 (r ~ 0.65) which indicates that 41.6% of the variance in the final grade can be explained by those predictor variables. In other words, the final score of a student in these three courses is strongly related to his/her prior learning performance (high school average score, GPA), English language skills, completion status at the university (at risk level, retention status, accumulated credit points) and participation in student support services (number of advisement received). The probability that this model could be developed using a different dataset is approximately 95%.

Applying the model:

The model was applied to the sample data to produce 'focus lists" of students that were statistically more likely to struggle. These lists were added to the LMS (Blackboard) course shells as groups.

To further enhance the system, research was also conducted inside the courses regarding LMS behaviour and early assessment results to ascertain correlations to final course grades. The table below outlines the results:

Internal course triggers from the LMS

Course		Trigger 1	Trigger 2	Trigger 3	Trigger 4
Introduction to Management	Trigger type	Quiz Attempts	Online Test	Individual Paper	Online test + Individual Paper
	Correlation	r= .329	r= .640	r= .843	r= .935
	Threshold	<1	<18	<20	<42
	Correlated to final score	>60	>60	>60	>60
Business Computing	Trigger type	Quiz Attempts	Assessment 1.1 + 1.2	Assessment 1.1 +1.2+2	Assessment 2
	Correlation	N/A	r= .653	r=.843	r= .671
	Threshold	Less than 2 attempts	<41	<70	<26
	Correlated to final score	N/A	<60	<60	<60
Marketing Principles	Trigger type	BB Activity	Assessment 1	Assessment 2	
	Correlation	r= .368	r= .407	r= .558	
	Threshold	<1hr 70% below the average	<=16	<=16	
	Correlated to final score	<61	<60	<60	

Based on the findings above, columns were added to the course shell Grade Centres. This then allowed the use of blackboard's Smart Lists, which permit the user to create lists based on multiple criteria, to create a dynamic "focus list" of student's who were not only identified by the statistical model, but

Evaluation of the model:

The model in the pilot was quite accurate. It predicted 63.22% of fails across the three courses and predicted 71% of all scores within a 10% error.

Week 0

Statistical model applied to current enrolments

Week 1

Teaching teams select meaningful triggers based on LMS behaviour, early assessment, feedback mechanisms or teacher observations

Week 2

Focus lists distributed to relevent teachiing teams

Week 3

Focus lists uploaded to the LMS

Week 4

LMS configured to generate focus lists that combine the statistical prediction with dynamic data from the LMS and/or the classroom

Week 5

Students on focus lists contacted by their teacher and invited to a meeting

Week 6

Teaching teams meet with students and refer to academic support services where appropriate

Week 7

Teaching teams feedback on the system:

All first year courses, Semester 1, 2017

The first year model:

The same supervised machine learning process was applied to all first year course data. Unsurprisingly, the first year model generated almost exactly the same predictor variables as the pilot with one exception. For this model, gender became an indicator, with male students scores being one grade point less compared to females if all other variables are equal. The R-squared of the first year model was slightly better than the pilot at 0.44, explaining 44% of the variance in final grades.

Accuracy:

At the time of this poster being created, final grades for this semester had not been released so the final assessment of the accuracy of the model is still pending. However, with application of the model against the testing data, the model predicted 73% of fails and accurately predicted 70% of final grades within a 10% error margin.

List distribution and information sessions:

Information sessions for all teaching staff were held in weeks 2 and 3 of the semester. The sessions covered the background of the project, the research findings, and instructions on how to use the system and give feedback to the researchers.

Assistance in internal course research and setting up LMS triggers and Smart Lists offered to all faculty

.

Hard copy 'focus lists' distributed to teaching teams

Focus lists in the LMS:

Focus lists uploaded to Blackboard as a flagged column in Grade Centre. This was chosen as the most convenient method as it allows easy viewing by faculty and the integration with Smart Lists.

Teacher intervention:

Teaching teams were asked to try to make contact with students who meet both criterion of being on the statistically generated list and showing other indicators of struggle through their LMS behaviour, early assessment results or classroom activity. It was recommended that teachers invite these students to a teacher-student meeting and when appropriate, refer that student to Student Academic Success (SAS)- the university's academic support team.

Feedback:

Teaching teams were asked to provide feedback on the system through ticking three boxes on the focus list:
1. Was there a need to contact the student based on their performance or behaviour in the course?
2. Did the teacher meet with the student?
3. Did the teacher refer the student to academic support?

Conclusion:

The statistical model is proving to be incredibly accurate, predicting 71% of fails in the pilot and 73% in first year testing. Unsurprisingly, the most significant predictor variables mostly relate to prior academic performance. Interesting, in the context of an English university in Vietnam, English proficiency also becomes an important variable. The statistical prediction alone provides a powerful tool for teaching staff to have a greater awareness of who may experience difficulties in their course. However, it is only a tool, and needs to be integrated with information coming from within the course of study to be meaningful. Their still remains an important challenge for this project. More feedback is needed from teaching teams on whether teacher intervention took place (the uptake on feedback on the pilot was close to 0%). Without this information, a true assessment of the system cannot be undertaken.

Future Research:

The future of this project is twofold. Further research is needed into other potential predictors, particularly in regards to geographical factors such as where the student went to high school (rural or urban) and the distance they live from campus while studying. Other areas of potential investigation include the educational history of their parents and socioeconomic status. After consolidating the research and the model, the project aims to dramatically improve on the existing university at-risk system and ultimately be integrated as a flag into the Student Record System. This could lead to a significant impact on student achievement, retention and a positive influence on teaching practice

STUDENT ACADEMIC SUCCESS

RMIT UNIVERSITY