Enhancing Student Success by Combining Pre-enrollment Risk Prediction with Academic Analytics Data

Dr. D. Raj Raman, Iowa State University

Raj Raman is Professor in the Agricultural and Biosystems Engineering (ABE) Department at Iowa State University, where he is also University Education Program Director and Testbed Champion for the NSF Engineering Research Center for Biorenewable Chemicals (CBiRC), Director of Graduate Education for the Interdepartmental Graduate Minor in Biorenewable Chemicals, and Education Programs Co-Leader for the USDA-AFRI project CenUSA Sustainable Production and Distribution of Biofuels for the Central USA. He is a licensed Professional Engineer who earned his BS in Electrical Engineering from the Rochester Institute of Technology and his PhD in Agricultural and Biological Engineering from Cornell University. Prior to coming to Iowa State in 2006, he was a faculty member at the University of Tennessee for over twelve years.

Raman enjoys teaching and has taught courses including freshmen engineering (mechanics and computer programming – to classes ranging in size from 20 to 500+), sophomore and junior level courses on mass and energy balance applications to biological systems engineering, numerical methods, electric power and electronics for technology students, senior design, as well as a long-standing residential/online graduate course on the fundamentals of biorenewable resources and technology. He has leveraged this interest into over $10M in teaching-related grant funding over his career and has contributed broadly to the literature in areas of curriculum, student risk characterization, and mentoring. He believes well trained, curious, thoughtful people are crucial to a university's research effort, and similarly to the function and survival of society. For this reason, the overarching goal of his teaching is to impart the core content needed by the students, and to do so while encouraging inquisition and higher levels of thought. He has secured competitive funds to support his teaching efforts – from university, industry, and federal sources – and for his efforts has received departmental, college, and national teaching honors including the Farrall Young Educator Award (2004) and the Massey-Ferguson Gold Medal Teaching Award (2016) given by the American Society of Agricultural Engineers. He has also been an invited participant in the National Academy of Engineering's 2013 Frontiers in Engineering Education Conference.

Raman chairs the ABE Engineering Curriculum Committee and in that role oversaw the successful 2012 ABET accreditation visit for both the Agricultural Engineering (AE) and Biological Systems Engineering (BSE) degree programs. Upon arriving at ISU in 2006, he led the development of the BSE program, and this program now enrolls over 100 students. Raman also runs multiple summer research internship programs through his roles in CBiRC and CenUSA – over 200 students have participated in summer programs he directed over the past decade. In his role as Pyrone Testbed Champion for CBiRC, Raman and his students have developed early-stage technoeconomic models of bioprocessing systems. His graduate students have gone on to faculty positions at peer institutions, and to engineering leadership positions at companies including Cargill, Nestle, and Merck.

Dr. Amy L. Kaleita, Iowa State University

Amy L. Kaleita is Associate Professor of Agricultural and Biosystems Engineering at Iowa State University, and a licensed professional engineer. She has a B.S. in Agricultural Engineering from Penn State University, an M.S. in Civil and Environmental Engineering from the University of Illinois at Urbana-Champaign, from which she also has a PhD in Agricultural Engineering. Her disciplinary research is in the area of data mining and information technologies for precision soil and water conservation.

Enhancing student success by combining pre-enrollment risk prediction with academic analytics data

Abstract
For nearly a decade, our institution has used multiple-linear-regressions models to predict student success campus-wide. Over the past three years, we worked to refine the success prediction models to the college of engineering (COE) students in particular, and to explore the use of classification and regression tree (CART) approaches to doing the prediction (Kaleita et al., 2016). In a parallel effort, our institution has contracted with an academic analytics company to do a university-wide retrospective analysis of course-level student performance in relation to graduation rates. Here, we report on recent work we have done to make synergistic use of the results from the COE CART model and the academic analytics. Specifically, we have been able to examine student performance (i.e., grades) in core "success marker" courses as a function of the risk-grouping into which the CART model places them. We are now using this information to inform our advising. We provide details on these efforts, and on the opportunities and challenges provided by data-driven approaches to enhancing student success.

Introduction
The academic success of students who enroll in engineering degree programs is critical to both the individual student due to the impact on earning potential (Carnevale et al. 2014), and to society. While it is recognized that a constellation of factors lead to attrition from engineering degree programs (Geisinger and Raman, 2013), encouraging students to take up a course of study for which they are not prepared is unethical, as we have argued elsewhere (Kaleita et al., 2016).

In prior work (Kaleita et al., 2016), we examined approaches to predicting the success of students in a large engineering degree program, based upon readily accessible pre-enrollment data. We showed that a classification and regression tree (CART) approach could correctly identify "at risk" (i.e., below 2.0 first-term GPA) students slightly more reliably than multiple linear regression. We also found that high school GPA was the strongest predictor of success, with standardized test (ACT) scores also highly ranked in predictive ability. Furthermore, although the model was built using 1st-term GPA as the objective function, we showed that applying the model to an incoming class from several years prior, and examining actual graduation rates, confirmed that the various risk categories had significantly different graduation rates, both within and outside engineering.

In a parallel and unpublished effort, our institution contracted with an academic analytics firm (EAB, a subdivision of *The Advisory Board Company*, Washington, DC) to undertake a longitudinal analysis of student success at our institution, with a focus on how examining grades in courses reflect graduation rates by university, college, or major. The results from the EAB effort allowed individual degree programs to evaluate the linkages between course grades and student graduation rates. Furthermore, it allowed establishment of success thresholds in key classes based upon a desired graduation rate.

The motivation for the work reported herein was to combine our risk-prediction efforts – which allowed grouping students into multiple risk categories based upon a small set of pre-enrollment data – with our insights into key success predictor courses. Specifically: What grades, on average, do different risk category students earn, in each of our key success prediction courses? A related question was as follows: To what degree are students in each risk category achieving the success threshold grade in the key courses? We believed that answering these questions might help us better support at-risk students in multiple ways, including (1) providing data-based advice on which courses to focus on; (2) providing insight into how course scheduling might impact performance (by comparing first vs. second semester performance in key courses); (3) helping advisers identify students who were missing success thresholds in key courses. Ultimately, we hoped that we could improve student success by these efforts.

Materials and Methods

As we have previously summarized (Kaleita et al., 2016), classification and regression tree approaches use machine-learning to build a rule set which predicts the output class based on a range of input variables (e.g., Margineantu and Dietterich, 1999). There are several advantages of CART models over regression models include that at-risk identification is inherently a classification problem, in which any student is predicted to be either at-risk or not. In addition, CART approaches are better suited to handling nonlinearities (including categorical variables), non-monotonic responses of the independent variable to changes in dependent variables, and variable-to-variable interactions.

Furthermore, CART has the ability to account for different costs or values of Type I (false positive) and Type II (false negative) errors, through a user-specified “loss matrix” that accounts for this asymmetric misclassification cost ratios (Margineantu and Dietterich, 1999; Bradford et al., 1998). In a linear regression approach, the costs associated with Type I and II errors are implicitly equal. Regarding this last point, we believe the cost associated with a false negative (failing to identify an at-risk student) is higher than that associated with a false positive. The cost of not providing intervention to students that might benefit from it is difficult to quantify (Veenstra, 2009) and, to our knowledge, has not been investigated. In prior work (Kaleita et al., 2016), we explored cost ratios ranging from 1:1 (i.e., equal cost) to 10:1 using data from the classes of 2010 – 2012. Unsurprisingly, the 1:1 CART model produced results similar to those from the linear-regression approach. For the work described herein, an updated 5:1 CART model was generated by the Iowa State University Enrollment Research Team (ISU ERT), based upon data from the 2015 incoming class. This CART5 analysis resulted in seven risk groupings, as summarized in Table 1 (following page).

The new CART5 analysis contained one particular grouping (Risk Group 2, or R2), which represented a small (3%) segment of the population on the basis of a self-identified metric, namely their “intent to participate in extracurricular activities.” We considered R2 to be an artifact of us not having specified that the model only use direct measures, and we excluded this group from subsequent analyses, and do not present it in Table 1 either.

Table 1: Criterion leading to six key risk groupings, as well as distribution of students amongst groupings. HS GPA = High School Grade Point Average; COE = College of Engineering; ACT El. Alg. = ACT Elementary Algebra sub-score; R1 = Risk Group 1, R2 = Risk Group 2, etc.

- HS GPA > 3.7
    - R1 (49% COE population, 96% > 2.0 first semester)
- HS 3.3 < GPA < 3.7
    - ACT Math > 22, ACT El. Alg. > Average
        - R3 (4% COE population, 94% > 2.0 first semester)
        - *1.5x more likely < 2.0 than R1*
    - ACT Math > 22, ACT El. Alg. < Average, Accepted early admission
        - R4 (12% COE population, 87% > 2.0 first semester)
        - *3x more likely < 2.0 than R1*
    - ACT Math > 22, ACT El. Alg. < Average, Accepted admission **not early**
        - R5 (8% COE population, 77% > 2.0 first semester)
        - *6x more likely < 2.0 than R1*
    - ACT Math < 22
        - R6 (3% COE population, 64% > 2.0 first semester)
        - *9x more likely < 2.0 than R1*
- HS GPA < 3.3
    - R7 (22% COE population, 58% > 2.0 first semester)
    - *10x more likely < 2.0 than R1*

We used the following criteria to select "core courses" for study: (1) math, science, and COE-designated courses common to all of the 10+ curricula in the COE; (2) departmental courses equivalent to COE-offered courses (e.g., with a high degree of commonality with either a central course or multiple other departmental courses), plus (3) courses identified by the ISU ERT as being taken by a large fraction of COE freshmen. On this basis, a group of 26 core courses were selected and provided to the ISU ERT.

To protect student privacy, student names were replaced with non-traceable subject identification numbers by the ISU ERT. The ISU ERT then provided spreadsheets with raw data showing grades in courses by subject identification number, as well as risk categorization for each subject identification number. These files were shared via secure file storage systems.

With the raw, de-identified data in hand, we computed average course grade point earned (CGP) for each risk category (based on both fall and spring grades). When we did this, some unsurprising trends in average grade structure appeared – that is, expected "easy" courses had high grades for virtually all risk groups, compared to well-known "hard" or "weed out" classes that did not. It appeared illustrative to rank order the courses in way that would reflect CGP of all groups. To do this, we fit a straight line to the CGP vs. Risk Group data, and computed the predicted CGP of a mid-risk student. We then rank-ordered the courses from highest to lowest predicted grade for a mid-risk student.

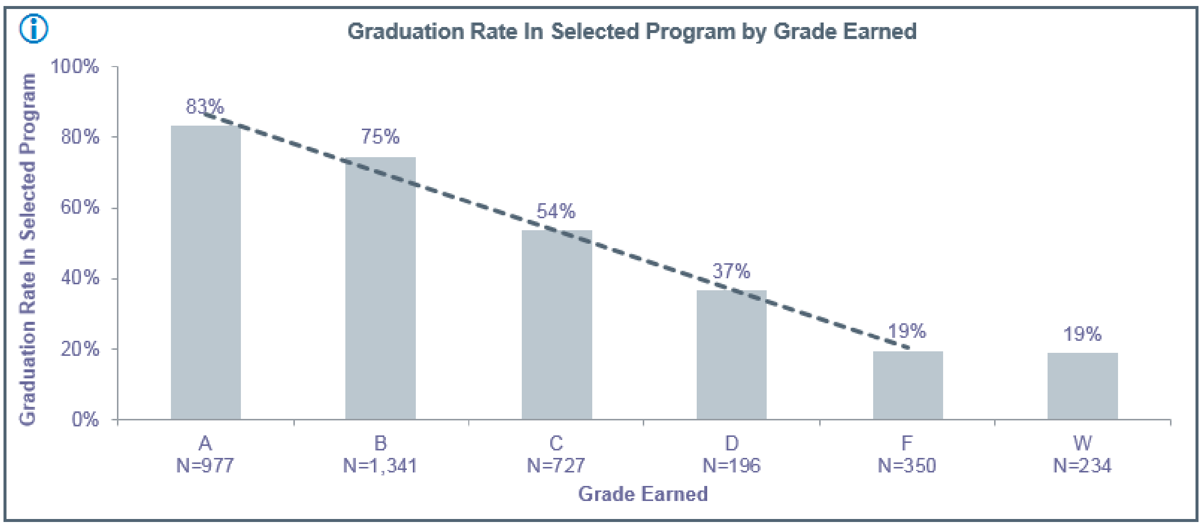
Once the 26 courses had been rank ordered from easy to difficult, we wondered if the ranking might somehow be used to create a *strength-of-schedule* score for a particular student. That is, might we characterize a given grouping of courses as "hard" or "easy" along some quantitative axis? We elected to take the following approach to compute such a metric: First, assign each course a strength rating (s_i represents strength of course i), according to the inverse of the CGP earned by risk-category 4 students. Now, compute strength-of-schedule as follows:

$$Strength\ of\ Schedule = \sum s_i a_i + \left(Total\ Eng\ Credits - \sum a_i\right) s_{min}$$

Where a_i = the credit hours associated with the i^{th} course, and where the term right of the plus sign effectively corrects for the credits taken in non-classified courses (using an assumption that they are relatively easier). After creating this method of assigning a strength-of-schedule for each anonymous student represented in the data, we were able to examine the distribution of GPA as a function of strength-of-schedule by risk grouping.

A recent longitudinal analysis of student success at our institution insight into graduation rates (from specific major, or college, or university) by grade cohort in any class. An example of the results of this analysis is shown in Figure 1, with the major and course removed for privacy reasons.

Figure 1: Sample data from external analysis of our program. This de-identified chart shows the graduation rate vs. grade earned in one particular program, for one particular class.



The data shown in Figure 1 was available for the core courses we selected. Furthermore, our own department (not the entire college) reviewed our core courses (a subset of those used in the overall work), and established "success marker" thresholds based upon a target 50% graduation rate. This provided a unique opportunity to further contextualize the grades earned by various risk groups, to see how various risk groups achieved, or missed, the success marker thresholds. We conducted and report on that portion of the analysis strictly for our own degree program.

Results and Discussion

When the student performance was examined by risk group and core course (Table 2), the core courses were found to have total enrollments (over both semesters) ranging from 49 to over 2700, with a mean of 740 and a median of 400. The performance of students in the various risk categories varied greatly across and within the 26 courses, as is easily seen in Table 2. For example, several core departmental engineering courses near the top of the table had midpoint students (R4) earning 3.4 CGPs, while these same risk-group students earned a full grade lower on average in science, engineering, and math courses near the bottom of the table. Similarly, differences between R1 (least at-risk) and R7 (most at-risk) grades varied from slightly over 1.4 grade points (more than the difference between a B- and an A at our institution), to essentially zero for a large science course that enrolls many non-engineering students. It seems noteworthy that even some of the "easier" courses found on the upper half of the listing have fairly significant gaps between the grades earned by R1 and R7 students.

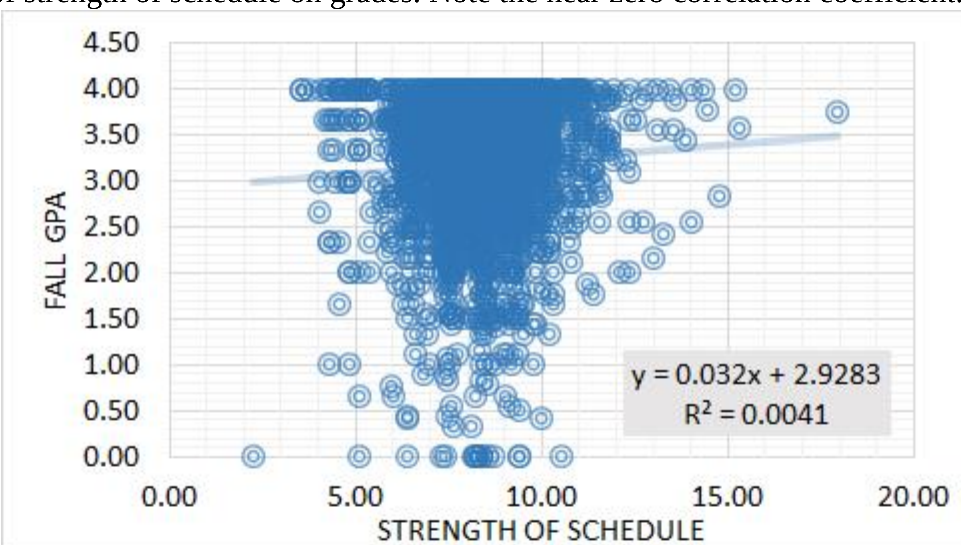
Table 2: Average course grade point earned (CGP) and standard deviation (4 point scale) by risk groupings for 26 core courses. Ordered by highest to lowest projected grade earned by R4 grouping. Color: yel = core eng course; blu = core sci.; grn = core math course; brn = other course; lt red < 2.50.

Course ID	n	R1	R3	R4	R5	R6	R7
ENGD 1xx	158	3.72±0.33	4.00±0.00	3.43±0.86	3.31±1.01	3.56±0.32	2.98±1.07
SCI 1xx	71	3.43±0.59	4.00±0.00	3.53±0.53	3.20±0.27	3.33±0.47	2.75±0.86
ENGD 1xx	182	3.75±0.51	3.79±0.48	3.72±0.57	3.21±0.83	2.22±1.6	3.53±0.75
ENGD 1xx	218	3.81±0.62	4.00±0.00	3.48±0.85	3.45±0.96	2.59±0.98	2.84±1.21
ENGD 1xx	808	3.64±0.55	3.08±0.96	3.44±0.58	3.22±0.93	3.18±0.77	2.91±0.98
SCI 1xx	504	3.09±1.00	2.77±0.76	3.20±0.82	3.10±0.94	3.44±0.35	3.16±1.02
OTH 1xx	776	3.26±0.62	3.25±0.62	3.05±0.72	3.03±0.66	2.95±0.85	2.64±0.97
SCI 2xx	49	3.72±0.57		3.26±0.13	1.84±0.17		3.06±0.71
MATH 1xx	119	3.31±0.52	4.00±0.00	2.75±0.84	3.29±0.6	2.53±1.11	2.17±1.21
ENGD 1xx	782	3.43±0.68	3.00±1.04	3.04±0.79	2.78±1.11	2.67±1.24	2.32±1.13
ENGD 1xx	413	3.53±0.68	2.88±0.97	3.35±0.65	2.72±1.14	2.29±1.23	2.44±1.26
ENGD 1xx	335	3.41±0.85	3.13±0.99	3.00±0.89	2.51±1.04	2.50±0.50	2.53±1.27
OTH 1xx	1822	3.41±0.78	3.13±0.8	2.97±0.89	2.85±0.93	2.28±0.84	2.40±1.12
ENGD 1xx	309	3.39±0.90	3.41±0.97	3.13±1.03	2.46±1.31	2.17±1.27	2.13±1.27
ENGD 1xx	569	3.27±1.01	3.37±0.88	2.89±0.99	2.75±1.17	1.77±1.11	2.17±0.71
ENGD 1xx	152	3.19±0.82	3.00±0.00	2.98±0.91	3.00±0.71	1.86±0.61	1.92±1.09
MATH 1xx	2114	3.06±0.91	2.61±1.12	2.46±1.04	2.60±0.99	2.37±0.97	2.17±1.21
MATH 1xx	393	3.00±0.68	3.56±0.42	2.34±0.9	2.60±0.90	1.88±0.81	2.08±0.97
ENGD 1xx	168	3.11±0.79	3.44±0.42	2.52±1.01	2.22±0.96	2.09±1.21	2.06±1.16
SCI 1xx	2408	3.14±0.77	3.07±0.81	2.54±0.89	2.59±0.95	1.76±0.96	2.18±1.01
SCI 1xx	1160	3.17±0.82	3.14±0.74	2.49±0.84	2.34±1.05	1.64±1.04	1.87±1.19
ENGD 1xx	623	3.26±0.89	3.04±1.01	2.46±1.10	2.26±1.20	1.33±0.00	2.23±1.23
SCI 2xx	81	3.00±0.94		2.70±0.83	1.96±1.07	2.33±0.47	1.57±1.21
ENGD 1xx	309	3.00±0.90	2.47±0.96	2.24±1.00	1.90±1.06	2.67±0.72	1.59±1.09
MATH 1xx	2745	2.98±0.96	2.75±0.92	2.35±1.05	2.33±1.16	1.47±1.05	1.88±1.14
SCI 2xx	1939	2.86±1.03	2.49±1.14	2.22±1.09	2.26±1.14	1.60±1.02	1.92±1.23

Using the analytics data, we identified six success marker courses: two core science classes (physics I and chemistry); two core engineering courses (engineering problem solving; engineering graphics and introductory design), one key math class along with a proxy for less prepared students – with a different threshold level (calculus I or pre-calculus) ; and one "other" category humanities class (composition). When we overlaid the success marker thresholds for our program on the results shown in Table 2, we found that risk groups 1 and 3 (the two lowest risk groups), were on average achieving all six thresholds. Risk groups 4 and 5 were achieving five of six thresholds, with a notable gap in one of the core engineering courses. But the two highest risk groups – R6 and R7 – were on average achieving only one or zero of the markers. Seeing this result is making us think carefully about how best to guide the freshmen in that category, who make up a quarter of the incoming class at our institution.

The impact of strength-of-schedule on student GPA was found to be negligible, despite our multiple attempts to examine the data in slightly different ways. Figure 2 (following page) is an example of the results for risk groups 1 and 3 (the two lowest risk categories) – but the scattershot nature of the plots were the same for higher risk groups, although the range of grades shifted as expected. In short, easier schedules were not observed to result in better average GPAs for fall or spring, for any risk group.

Figure 2: Strength of schedule data for R1 and R3 combined, illustrating the mild *positive* influence of strength of schedule on grades. Note the near zero correlation coefficient.



We examined differences in fall and spring grade distributions by class and risk group (data not shown), and discovered that there is an almost universal sag in GPA during the second semester. We had expected that perhaps some of the higher risk groups would do better in harder classes when taken in the second semester, but no quantitative evidence for this expectation was found.

Conclusions

Categorizing the at-risk status of students in a large college of engineering provides a unique lens through which to examine student performance in subsequent classes. Some of our expectations – for example regarding the value of delaying a difficult required class, or of taking a light schedule – are not supported by such analysis. But overlaying a knowledge of key success marker thresholds on this approach shows that some risk categories are predicted to *systematically miss these markers*. Multiple approaches to operationalizing these data – ranging from asking advisers to check on mid-semester performance of high-risk students in success marker classes, or by offering additional instructional support for success marker classes and doing targeted advertising to high-risk students, or to changing admission thresholds to exclude the highest risk students from a program where they are unlikely to succeed – are conceivable. We have taken small steps in this regard – sharing with students the nature of our work and providing guidance on success strategies, sharing with administration our findings and doing follow-on work aimed at understanding the risk-categorizations of our recent graduates – but more work is needed to help increase the graduation rates of all students entering our program.

References

1. Kaleita, A. L., G. R. Forbes, E. Ralston, J. Compton, D. Wohlgemuth, and D. R. Raman. 2016. Pre-enrollment identification of at-risk students in a large engineering college. *International Journal of Engineering Education* 32(4): 1647 – 1659
2. Carnevale, A. P., J. Strohl, and M. Melton. 2014. *What's It Worth? The Economic Value of College Majors.* Georgetown University, Center on Education and the Workforce.
3. Geisinger, B. N., and D. R. Raman. 2013. Why they leave: Understanding student attrition from engineering majors. *International Journal of Engineering Education* 29(4): 914 – 925
4. Margineantu, D. D., and T. G. Dietterich. 1999. *Learning decision trees for loss minimization in multi-class problems*. Technical Report 99-30-03, Department of Computer Science, Oregon State University, Corvallis OR.
5. Bradford, J. P., C. Kunz, R. Kohavi, C. Brunk, and C. E. Brody. 1998. *Pruning decision trees with misclassification costs*. In Machine Learning: ECML-98, Proceedings of the 10th European Conference on Machine Learning, Chemnitz, Germany, April 21-23, 1998, 131 – 136.
6. Veenstra, C. P. 2009. A strategy for improving freshman college retention. *Journal for Quality and Participation,* January 2009, 31(4):19 – 23.