

AUTHORS

Mark David Milliron, Ph.D.
Civitas Learning

Laura Malcolm, M.A.
Civitas Learning

David Kil, M.S.E.E., M.B.A.
Civitas Learning

Abstract

Civitas Learning was conceived as a community of practice, bringing together forward-thinking leaders from diverse higher education institutions to leverage insight and action analytics in their ongoing efforts to help students learn well and finish strong. We define insight and action analytics as drawing, federating, and analyzing data from different sources (e.g., ERP, LMS, CRM) at a given institution to produce deep predictive flow models of student progression and completion coupled with applications (apps) that take these data and bring them to advisors, students, faculty, and administrators in highly consumable/useable ways. Through three case studies, this article provides a closer look at this iterative work unfolding in diverse institutions, addressing diverse student success challenges, and achieving significant positive results on student progression. The article underscores a key finding: there is not a one-size-fits-all predictive model for higher education institutions. We conclude with a discussion of key findings from these cases and observations to inform future related work.

Insight and Action Analytics: Three Case Studies to Consider

Civitas Learning was conceived as a community of practice, bringing together forward-thinking leaders from diverse higher education institutions to leverage insight and action analytics in their ongoing efforts to help students learn well and finish strong (Fain, 2014; Thornburgh & Milliron, 2013). Our fast-growing community of practice now includes more than 40 institutions and systems, representing more than 570 campuses, serving more than 1.45 million active students. It includes research one institutions, emerging research and access universities, independent colleges, community colleges, and private sector universities. We work with cross-functional groups of administrators, IT teams, IR teams, advisors, and faculty members, most of whom are leading large-scale student learning and completion programs, often catalyzed by federal, state, foundation, and institutional dollars, pressures, and aspirations. Some initiatives include the Obama Administration 2020 Goals (Higher Education, 2014), Complete College America (2014), Bill & Melinda Gates Foundation Postsecondary Initiative (Postsecondary success strategy overview, 2014), Lumina Foundation for Education's Goal 2025 (Lumina Foundation Strategic Plan, 2013), Texas Student Success Council (2014), Hewlett Foundation's Deeper Learning Initiative (2014), and Kresge Foundation's Education Initiative (2014; Milliron & Rhodes, 2014). It is important to note that we do not conceive of our work as another new initiative. Indeed, many of these institutions report that they are already reeling from "initiative fatigue." Rather, our insight and action analytics infrastructure is meant to be a powerful resource to try, test, and power deeper learning and student success initiatives (Kim, 2014).

CORRESPONDENCE

Email
mark.milliron@civitaslearning.com

We define insight analytics as the family of activities that bring data from disparate sources together to help create a more complete view of student progression. In the most basic terms, this means (a) federating data from an institution's Student Information System (SIS) and Learning Management System (LMS); (b) using sophisticated data science tools and techniques, including machine learning, data availability segmentation and clustering, to create and compete feature variables derived from the diverse sources; (c) building an array of predictive models; and then (d) leveraging a variety of visualization techniques we explore the resulting historic and predictive student progression/flow

models for insights that help better understand how students succeed and face challenges on their higher education journeys. Once the models are developed, we create a cloud-based, production-quality, predictive-flow-model infrastructure for each institution that is updated at minimum on a rolling five-term cadence to keep the student-level predictions as current as possible. From here, more sophisticated insight analytics work includes adding additional data sources in this mix, such as Census, application data, card swipe, CRM, and more, and then testing these new data streams for added predictive power to drive decisions about how or whether to add them to the production system. See the Appendix for a deep dive on some of these techniques.

We created a platform application called Illume™ that brings insights from this work to our institutional partners, allowing them to view student progression dynamics filtered by chosen segments (e.g., part-time, full-time, Pell recipients, distinct campuses, members of intervention category), often testing assumptions about performance and possible historic and predictive trends (Figure 1.1). The application also surfaces powerful predictors for distinct segments, which are feature and point variables contributing significantly to the success or challenge of a given segment. For example, a feature variable we derive called affordability gap – the delta between financial aid received and tuition owed – is often a far more powerful predictor for first-time students than placement test scores. The diverse segment and cluster analyses often point to relationships that are non-intuitive or surprising, and other times reaffirm long-held assumptions. Either way, they are useful in starting conversations about tipping points, momentum points, and possible dynamics at work in systems, processes, policy, and practice at the institution.

It is important to note that we do not conceive of our work as another new initiative. Indeed, many of these institutions report that they are already reeling from “initiative fatigue.”

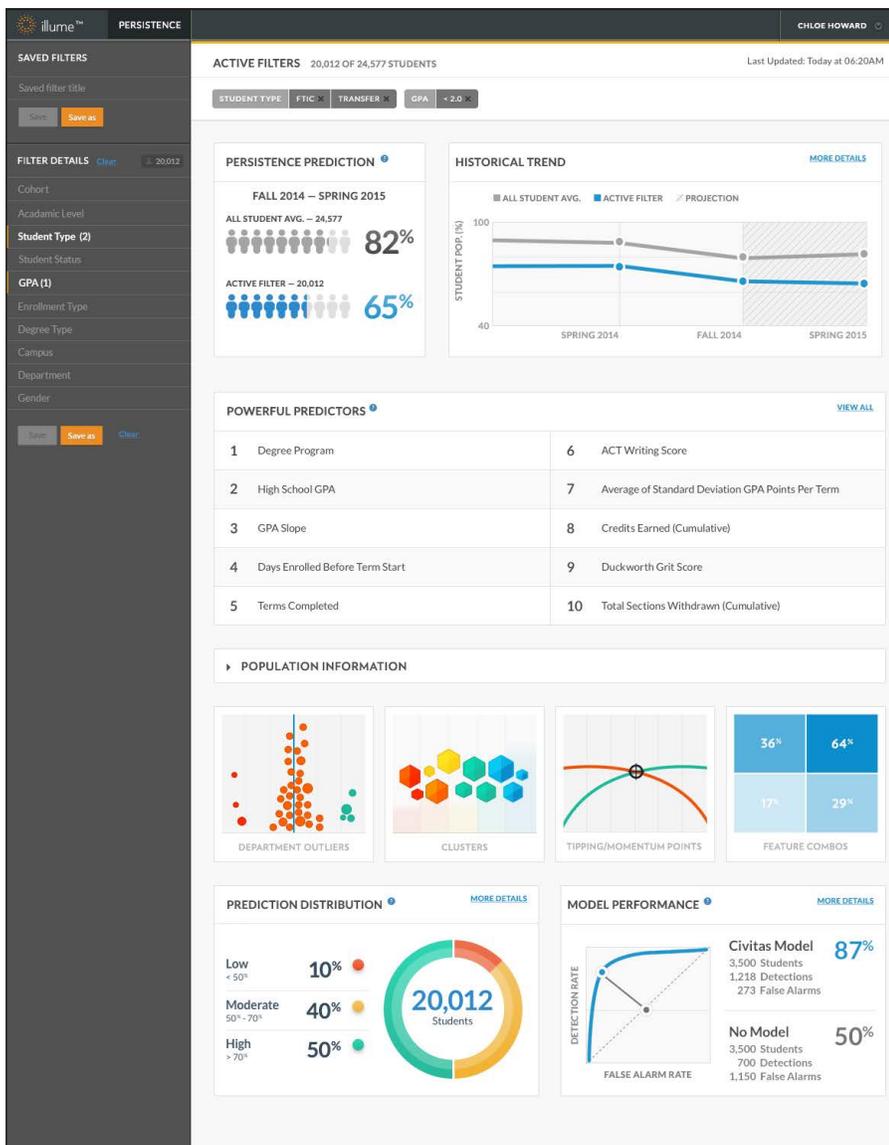


Figure 1.1

This insight analytics infrastructure can be useful, to be sure. But in our work over the last three years we have found that this predictive flow platform is more a predicate than a solution. The insights derived can make a stronger impact on student success when used to power action analytics. Action analytics include applications (apps) that use design thinking, user–interface disciplines, and strategic workflow to leverage insight analytics in easy to consume, engaging, and highly useable formats to help administrators, advisors, faculty, and even students interact with these data to help understand risk and success factors, target and test interventions, and guide choices, outreach, and activity. We have developed a family of action–analytic apps that include our Degree Map™, Inspire™, and Hoot.Me™ family of apps (Figure 1.2). Each of these is being deployed at different institutions and are being tried, tested, and tuned as the work of learning about how to bring insight and action analytics into the daily operations of institutions continues.

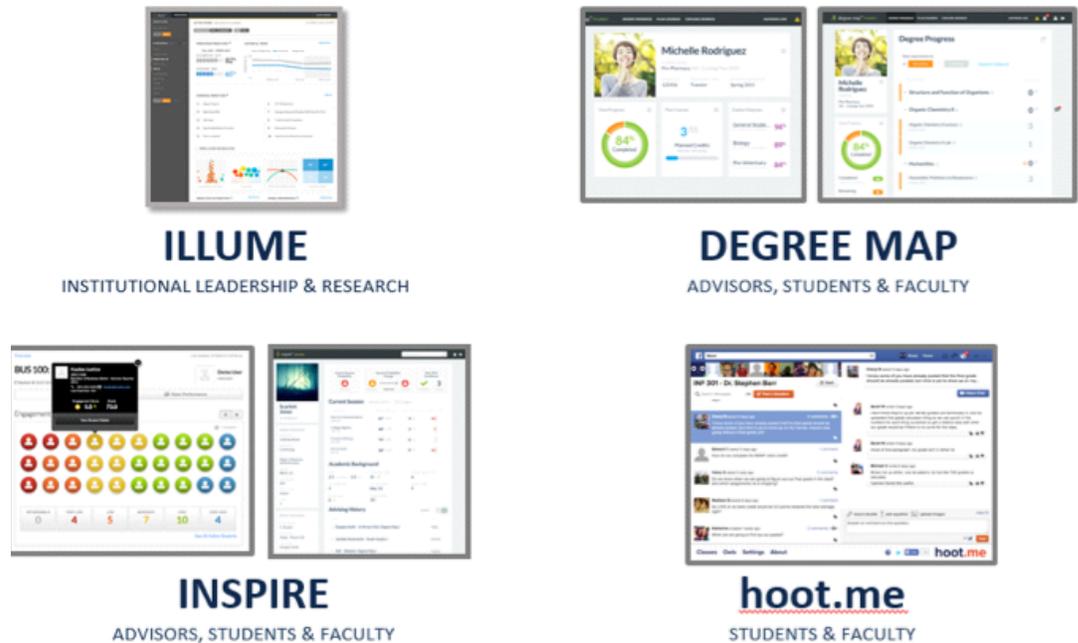


Figure 1.2

Rather, our insight and action analytics infrastructure is meant to be a powerful resource to try, test, and power deeper learning and student success initiatives.

There is, of course, an array of learning–centered and student–completion–centered action applications at work in the field of higher education, from basic early–alert systems to comprehensive CRM tools (Blumenstyk, 2014; Milliron, 2013). However, most of these have choice architectures and engagement tools powered by heuristic triggers and set configurations as opposed to institution–specific, student–level predictive flow models. Others leverage quite sophisticated advanced analytics, but only in the context of their application (e.g., several adaptive learning tools). However, many of these action applications are likely to add insight–analytic linkages on the road ahead and will move into a growing ecosystem of what we call Action Analytic Applications. Indeed, we are likely to see dozens, if not hundreds of these, emerge in the months and years ahead.

It is important to note that these action analytic applications can be data streams in and of themselves that can inform and improve the insight analytics work, creating an ongoing and continuously improving analytics infrastructure. For example, both the Inspire for Advisors and Inspire for Faculty Apps generate data on tried interventions with different students that can inform future suggestions for advisors and faculty members. Hoot.me, which is a crowd–sourced, student–driven, question–and–answer community app generates engagement and social interaction data. Indeed, some future action analytic application may be used primarily to generate data – e.g., an app that gathers wellness behaviors or non–cognitive mindsets through micro surveys.

The interplay between and the process of learning more about insight and action analytics has been at the heart of our work for the last three years. The community of practice

site, Civitas Learning Space, showcases the ongoing initiatives in an effort to inform and engage a broader audience. Moreover, the Civitas Learning partner community comes together twice a year for summits on data science, intervention strategies, and future planning (Rees, 2014).

What follows is a closer look at three of our partner institutions as they brought together their insight and action analytics initiatives. We present three cases in an effort to show how this iterative work unfolds in diverse institutions, approaching diverse student success challenges, and to underscore a key finding: There is not a one-size-fits-all predictive model for higher education institutions. Each institution has its own predictive student flow and leaders, teachers, and advisors need to understand and engage their student success strategies in the context of their own students, policies and practices. We will come back in the concluding section to offer observations for those interested in learning more or joining in similar efforts.

Case Study One: Early Intervention for Course Success

Executive Summary

Leveraging Civitas Learning's Illume predictive analytics platform and Inspire application for administrators and advisors, Partner Institution A ran a pilot program to test the efficacy of using predictive-analytics-based interventions on driving improvements to student course completion rates. Over the course of three terms starting in Spring 2013, predictive models were built, approaches to intervention were tested, and outcomes were evaluated using a randomized test and control pilot approach. In the first two terms of the pilot, no statistically significant improvements to outcomes were measured. In Fall of 2013 with a pilot group of ~14,000 enrollments (~7,000 each in test and control) and applying learnings from previous terms, the institution realized an average improvement of 3% at a 98% confidence level for statistical significance test vs. control. This translates into 210 student enrollments that successfully completed their course that otherwise would have failed or withdrawn.

We define insight analytics as the family of activities that bring data from disparate sources together to help create a more complete view of student progression

Introduction

Institution A is a 4-year access institution with greater than 50,000 students including undergraduate and graduate. They offer on-campus programs and courses as well as online programs through an online campus.

The focus of the pilot with Institution A was using advanced analytics to understand student risk, the variables that contribute most to student success, and most importantly how to make these insights actionable to improve student outcomes. Ultimately, the institution goal is a more personalized student experience and a better probability for student success, which translates to higher course completion, retention, and graduation rates to fulfill their institutional mission.

Methodology

Three pilots were conducted over the course of three terms (Spring 2013, Summer 2013, and Fall 2013) using randomized assignment of all enrollments within a section to test or control groups. While random assignment at the enrollment level would be preferred to reduce selection bias based on section and instructor, operational constraints prevented this approach.

In order to evaluate the potential section level bias, baseline predictions of course success were used to evaluate whether the sections were biased. Deltas between prediction of course success showed no statistically significant difference between test and control group in terms of enrollment likelihood to successfully complete.

In all three pilots course success was defined as finishing the course with a grade of C or better for undergraduate enrollments and B or better for graduate enrollments. For outcomes analysis, statistical significance was computed using Fisher's exact test, widely used in the analysis of contingency tables (Fisher, 1954).

In Spring 2013, nine courses (four graduate and five undergraduate) participated in the pilot with 2,279 enrollments in total. In Summer 2013, the pilot grew to ten courses

(five graduate and five undergraduate) and 6,832 enrollments. Finally, in Fall 2013 the pilot included 15,500 enrollments across 25 courses (10 graduate and 15 undergraduate).

Study

While predictive analytics have the potential for wide applicability across the student lifecycle, the starting point for this pilot focused where there could be concrete results that could be measured in a short amount of time – student successful course completion.

Pilot goals were:

- Demonstrate that predictive analytics in combination with targeted interventions can improve student outcomes.
- Evaluate which interventions produce better outcomes.
- Learn from the process and determine strategies to scale predictive analytics for personalized interventions.

For example, a feature variable we derive called affordability gap – the delta between financial aid received and tuition owed – is often a far more powerful predictor for first-time students than placement test scores. The diverse segment and cluster analyses... are useful in starting conversations about tipping points, momentum points, and possible dynamics at work in systems, processes, policy, and practice at the institution.

Pilot roll-out. Leveraging historical data, Civitas Learning developed institution specific predictive models to evaluate the complex set of variables contributing to student success. These models provide an individualized risk prediction of each student's likelihood to successfully complete a course, with greater than 80% accuracy prior to the course start. As student behaviors were introduced into the models over the course term, the student's risk prediction was continually updated, providing an increasingly accurate measure of course completion likelihood.

Civitas Learning's Inspire application delivered these predictions in an actionable way to academic administrators and advisors, so that they could understand which enrollments were at-risk and apply timely interventions and support. Users analyzed data, segmented student populations and implemented targeted communications directly from the application.

Spring 2013 pilot. In the initial Inspire for Administrators roll out in the Spring of 2013, based on insights from the application, subgroups were analyzed to determine variance in probability to succeed based on many predictive factors (including GPA, attendance patterns, grades, terms of enrollment, course credits and many more). Email communications were sent from the Inspire application by academic program administrators based on student risk factors. Content of the emails was determined by the program administrator and varied across programs. Fifty-one percent of enrollments received an email intervention with an average of 1.71 interventions per enrollment. The control group did not receive interventions.

Summer 2013 pilot. In the Summer of 2013, using the same predictive model, academic program administrators expanded the pilot to a larger number of courses and enrollments. Again, email communications were sent from the Inspire application by program administrators based on student risk factors. However, in this pilot, the test group was broken into four sub-groups to test varied outreach approaches including templated content and timing differences. Outreach approaches for each test group were developed by a committee of academic leads from across programs. In the Summer pilot, 54% of enrollments received an email intervention with an average of 1.36 interventions per enrollment. The control group did not receive interventions.

Fall 2013 pilot. Deployment and experimentation with selected interventions allowed for early testing of intervention approaches during the spring and summer terms. Processes were operationalized and refined, and best practices were established regarding the dissemination of interventions in preparation for the Fall 2013 term.

In Fall of 2013, Academic Program Administrators and Advisors used the app (Inspire for Administrators) to determine students most in need of intervention, then pulled from a prepared suite of intervention tools, messaging, emails, and calendar items to provide support in a timely, empathetic, appropriate way to students in the test group. The control group did not receive interventions.

Findings

Model performance. Looking retroactively at model performance, at an individual student level, predictive models were able to identify with 83% accuracy on the first day of a course the students who would successfully complete and by day seven the accuracy level moved to 86%. Model performance remained at these levels across the three pilots.

Outcome performance. In Spring of 2013, the test group outperformed the control group in successful course completion by 122 basis points. However, the p -value was 0.2677 not reaching statistical significance. Institution A found these results to be promising and developed a series of templated outreach plans to facilitate outreach for the next term.

In Summer 2013, there was no measurable impact on successful course completion. Theories as to why there was no improvement focused on the complexity of the intervention outreach plans and the user base of the application. Institution A decided to simplify the outreach approach for fall and to add advisors to the pilot to assist with student outreach.

In Fall of 2013, the test group of ~5,000 undergraduate students outperformed the control group in successful course completion by 300 basis points. This result had a p -value of 0.05 reaching statistical significance at a confidence level of 95%. There was no measurable improvement for graduate students.

Case Study Two: Early Intervention by Faculty for Persistence Gains Executive Summary

Leveraging Civitas Learning's Illume™ predictive analytics platform and Inspire for Faculty application, Partner Institution B ran a pilot program to test the efficacy of using predictive analytics based interventions to drive improvements in student persistence rates. Over the course of three terms starting in Fall of 2012, predictive models were built, an application was launched to facilitate faculty outreach, and outcomes were evaluated. A pilot was conducted across two terms beginning in the Winter 2013 term. During the pilot, faculty used a "heat map" of student engagement to identify and prioritize students for intervention outreach. In the first term of the pilot no statistically significant improvement to outcomes was measured. In the Spring Term of 2014 with a group of ~68,000 online enrollments and applying learnings from previous terms, the institution realized statistically significant persistence improvements.

Introduction

Institution B is a 4-year access institution with more than 20,000 students including both undergraduate and graduate programs. They offer on-ground programs as well as an online campus. The focus of the pilot with Institution B was to use advanced analytics to understand online student risk of successful course completion and persistence and use that understanding for the prioritization and differentiation of outreach by faculty.

Methodology

Two pilots were conducted over the course of two terms (Winter 2013 and Spring 2014). The first pilot focused on undergraduate online students in six high enrollment courses. In the first term, randomized assignment of students to test and control groups created the pilot group. In the second term, because of operational challenges in administering interventions to only test students, propensity score matching was used to identify a matching control group. This allowed for all online enrollments to be in the test group while identifying the control group from historical enrollments.

Propensity-score matching (PSM) is used in observational studies when there is no randomized control group. Simply put, PSM compresses salient features (x) of pilot participants into a single variable called propensity score. It then computes the propensity scores of non-participants using their attributes and finds matching cohorts, such that $p(z=1/x) = p(z=0/x)$, where z is the binary participation variable. This assures that the matching cohorts are statistically similar to the pilot group in x . As an extra security layer, top features (x) from the predictive models ($y = f(x)$) are used in PSM. This ensures that the created control group

We present three cases in an effort to show how this iterative work unfolds in diverse institutions, approaching diverse student success challenges, and to underscore a key finding: There is not a one-size-fits-all predictive model for higher education institutions. Each institution has its own predictive student flow and leaders, teachers, and advisors need to understand and engage their student success strategies in the context of their own students, policies and practices.

is virtually indistinguishable from the pilot group from an outcomes (y) perspective. That is, $p(y/x, z=1) = p(y/x, z=0)$.

In all three pilots, persistence was defined as re-enrolling in the next term and staying enrolled past the add-drop/census period in the following term. For outcomes analysis, statistical significance was computed using Fisher's exact test, widely used in the analysis of contingency tables (Fisher, 1954).

In Fall of 2013...the institution realized an average improvement of 3% at a 98% confidence level for statistical significance test vs. control. This translates into 210 student enrollments that successfully completed their course that otherwise would have failed or withdrawn.

Courses participating in the pilot grew to all online courses in the second term. The student enrollment count in the Winter term was approximately 15,000 (with 7,500 each in test and control). However, in the Spring 2013 term due to including all online enrollments the pilot grew to ~68,000 enrollments each in test and control groups.

Study

While predictive analytics has many applications, this pilot focused on leveraging faculty outreach to drive improvements to student persistence through effective outreach.

Pilot goals were:

- Demonstrate that predictive analytics, in combination with targeted interventions, can improve student outcomes.
- Focus faculty on improving student engagement in online courses
- Learn from the process and determine strategies to scale predictive analytics for personalized interventions.

Pilot roll-out. Leveraging historical data, Civitas Learning developed institution-specific predictive models to evaluate the complex set of variables contributing to student successful course completion and engagement in online courses. These models provided an individualized risk prediction of each student's likelihood to successfully complete the course. From this model the online course behaviors predictive of course success were identified and used to create a student engagement score. The engagement score was based on a zero to ten point scale and was relative – comparing engagement to all other enrollments taking the same course at the same time. The engagement score weighted behaviors based on their contribution to the predictive model.

Looking retroactively at model performance, at an individual student level, predictive models were able to identify with 83% accuracy on the first day of a course the students who would successfully complete and by day seven the accuracy level moved to 86%.

Civitas Learning's Inspire application then delivered the engagement score in an actionable way to online faculty, so that they could prioritize and differentiate intervention outreach to students. In addition to the engagement score, key information was included to help faculty understand why students were at risk so they could apply timely interventions and support. Using the application, faculty emailed students to drive increased online course engagement. All outreach was tracked so approaches and timing could be analyzed for effectiveness. In addition, since engagement scores were relative, faculty could monitor their section engagement in order to see how their students were doing on engagement compared to the whole.

Winter 2013 pilot. In the initial pilot, conducted during the Winter 2013 term, the predictive models generated a daily engagement score for each student in each section. Faculty used this score to assist in prioritizing outreach for students in the test group. The interface provided direct access to their assigned sections and students as well as the ability to segment students for outreach based on parameters such as current grade in course, engagement score, etc. In addition, the interface allowed faculty to track interventions and see a record of all emails sent to a student.

Finally, a tracking dashboard was deployed that allowed faculty to track week to week progress on engagement, successful course completion and continuation and compare that progress between their section and all other sections of the same course. Faculty used this prediction to assist in prioritizing re-enrollment and differentiating outreach for students in the test group. Faculty used their standard instructional process for control group sections.

Spring 2013 pilot. In the Spring 2013 term, the application was enhanced to allow faculty to "bulk" email students. Bulk email provided faculty the means to email the

same content to multiple students, with name personalization, in one action. In addition, “Recommended Outreach” was added to the interface to provide quick links to faculty to assist completion of the most common interventions. For example, one recommendation filtered “students with low engagement who haven’t had outreach in the past week” and let faculty email them in one click.

Findings

Model performance. Looking retroactively at model performance by reviewing engagement scores in comparison to final grades, the data show that the scores were highly reflective of successful course completion.

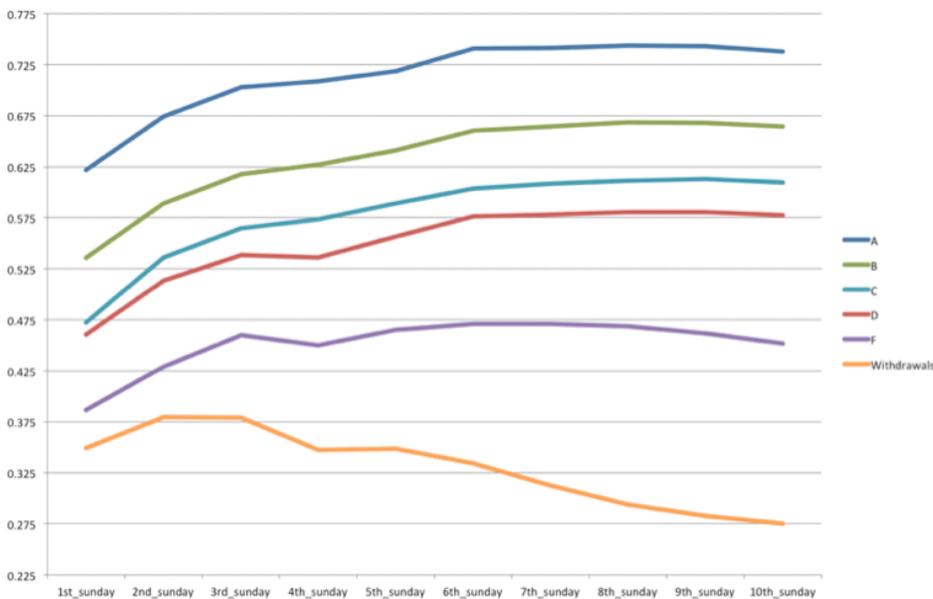


Figure 1.3. Illustration of the week by week engagement score trend in comparison with the student final grade shows the engagement score is highly correlated with successful course completion.

Outcome performance. In the Winter 2013 term, the test group outperformed the control group in persistence by 91 basis points. The result was not statistically significant reaching a p -value of 0.19 with a confidence level of 81%. However, institution B found these results to be promising and in the following term made plans to widen the pilot to include all online courses.

In Spring 2013, persistence rates from the Spring Term into the Summer Term were 321 basis points greater for test group than the control group. This result had a p -value of 0.05 reaching statistical significance at a confidence level of 95%. This result was calculated using retrospective propensity score matching to identify the control group. In order to validate the results a second analysis was done using time-series forecasting and the results held at a statistically significant level.

Case Study Three: Early Intervention by Advisors for Persistence Gains Executive Summary

Leveraging Civitas Learning’s Illume predictive analytics platform and Inspire application for Advisors, Partner Institution C ran a pilot program to test the efficacy of using predictive analytics based interventions on driving improvements to student persistence. Over the course of three terms, starting in January of 2014, predictive models were built, approaches to advisor led intervention were tested, and outcomes were evaluated using a randomized test and control pilot approach. In the first two terms of the pilot no statistically significant improvements to outcomes were measured. However, in the May 2014 term with a pilot group of ~10,000 students, and applying learnings from previous terms, the institution

In reviewing the intervention data by terms completed, for early term students, phone calls where the advisor spoke to the student were the most effective intervention. Conversely, for students with greater than ten terms completed at the institution, email appears to be the best initial intervention.

realized statistically significant improvements in persistence for students in their first nine terms. Largest gains were realized for new students, with a 762 basis point improvement in persistence when comparing the test to the control group.

Introduction

Institution C is a career-focused 4-year access institution with more than 40,000 students including both undergraduate and graduate programs. They offer on-ground campus locations as well as an online campus. The focus of the pilot with Institution C was to use advanced analytics to understand student risk of re-enrollment and persistence. In addition, the pilot was designed to use that understanding for the prioritization and differentiation of enrollment services through their advising function.

Methodology

Three pilots were conducted over the course of three terms (January 2014, March 2014, and May 2014). The pilot focused on undergraduate online students in six degree programs. In the first two terms, randomized assignment of students to test and control groups created the pilot cohort. In the third term, because of operational challenges in administering interventions to only test students, propensity score matching was used to identify a matching control group. This allowed for all students within the specified degree programs to be in the test group while identifying the control group from other degree programs.

Propensity-score matching (PSM) is used in observational studies when there is no randomized control group. PSM compresses salient features (x) of pilot participants into a single variable called propensity score. It then computes the propensity scores of non-participants using their attributes and finds matching cohorts, such that $p(z=1/x) = p(z=0/x)$, where z is the binary participation variable. This assures that the matching cohorts are statistically similar to the pilot group in x . As an extra security layer, top features (x) from the predictive models ($y = f(x)$) are used in PSM. This ensures that the created control group is virtually indistinguishable from the pilot group from an outcomes (y) perspective. That is, $p(y/x, z=1) = p(y/x, z=0)$.

In all three pilots, persistence was defined as re-enrolling in the next term and staying enrolled past the add-drop/census period in the following term. For outcomes analysis, statistical significance was computed using Fisher's exact test, widely used in the analysis of contingency tables (Fisher, 1954). Degree programs participating remained consistent across the three pilots. The student count in the January and March terms was approximately 5,000 (with 2,500 each in test and control). However, in the May 2014 term, due to including all students in the selected degree programs, the pilot grew to ~10,000 students with 5,000 each in the test and control groups.

Study

While predictive analytics has many applications, this pilot focused on using predictive analytics to maximize the effectiveness of advising resources in driving re-enrollment and student persistence.

Pilot goals were:

- Demonstrate that predictive analytics, in combination with targeted interventions, can improve student outcomes.
- Maximize application of advising resources to improve persistence.
- Evaluate which intervention approaches produce better outcomes and for which students.
- Learn from the process and determine strategies to scale predictive analytics for personalized interventions.

Pilot roll-out. Leveraging historical data, Civitas Learning developed institution-specific predictive models to evaluate the complex set of variables contributing to student persistence. These models provide an individualized risk prediction of each student's likelihood

Insight analytics that are developed using institution-specific data sources – particularly student-level SIS and LMS data – are vital to understanding student flow, as well as targeting and personalizing intervention and outreach.

In short, there is not a global predictive model that works across institutions with any level of accuracy. You need to “turn the lights on” in your institution.

to persist at the institution. As student behaviors were introduced into the models over the course term, the student's risk prediction continually updated, providing an increasingly accurate measure of persistence likelihood for advisors.

Civitas Learning's Inspire application delivered these predictions in an actionable way to advisors (student success coaches), so that they could prioritize and differentiate re-enrollment outreach to students. In addition to the prediction, key information was included to help advisors understand why students were at risk so they could apply timely interventions and support. Using the application, advisor managers analyzed data, designed outreach approaches, and assigned advisors to students for intervention. All outreach was tracked so it could be analyzed for effectiveness.

January 2014 pilot. In the initial pilot, conducted during the January 2014 term, institution-specific predictive models were used to generate a "Day 0" report that identified students' probability to persist into the following term starting the day before the new term. This model used student information system (SIS) data to make the prediction. Advisors used this prediction to assist in prioritizing re-enrollment and differentiating outreach for students in the test group. Advisors used their standard re-enrollment process for control group students.

A probability score between 0 and 1 was generated for each student and students were distributed into five persistence groups (quintiles) based on this score. Groups ranged from very high to very low probability to persist. Advisors were provided with the group assignment for each student along with key academic background information for context. Background information differed depending on whether students were new or continuing.

The report was delivered in the form of a spreadsheet to advisor managers who used it to make advisor assignments and design outreach approaches. Advisors used a combination of email and phone call outreach to test group students. Across the term, re-enrollment was tracked and reported to the group on a weekly basis.

March 2014 pilot. In the March 2014 term, the predictive models were enhanced to include learning management system (LMS) data. In addition, delivery of the spreadsheet moved from a one-time report to a report delivered nightly. As in the January pilot, advisors used this prediction to assist in prioritizing re-enrollment and differentiating outreach for students in the test group. Again, advisors used their standard re-enrollment process for control group students.

May 2014 pilot. In the May 2014 term, the report was replaced by the Inspire for Advisors application which provided a user interface for each advisor to manage their student caseload. The interface provided direct access to their assigned student list as well as the ability to segment students for outreach based on parameters such as degree program, new vs. continuing status, probability group, and recent changes to their probability. In addition, the interface allowed advisors to track interventions and see a record of all outreach administered to a student. Finally, a re-enrollment tracking dashboard was deployed that allowed advisor managers to track week to week progress on continuation and compare that progress between the test and control groups. As in the previous pilots, advisors used this prediction to assist in prioritizing re-enrollment and differentiating outreach for students in the test group. Again, advisors used their standard re-enrollment process for control group students.

Findings

Model performance. Looking retroactively at model performance by reviewing the probability group assignments, the data show that the predictions were highly reflective of actual student persistence rates. For example, for students in the 0–40% probability of persistence range, average actual persistence was 27%. On the other end of the spectrum, for students in the 80–100% probability of persistence range, average actual persistence was 86%.

Figure 1.4 shows the actual Receiver Operating Characteristic (ROC) curves for Institution C to explain salient concepts. Assuming the intervention outreach capacity of 10K students, using the purple model (test) provides 141% improvement (20.9% to 50.5%) in correctly identifying eventual non-persisting students for intervention in comparison to randomly reaching out to students.

Bringing insight analytics together with action analytics is essential to "moving the needle" on student success. Better precision of the models helps target outreach and improve impact of instruction and advising support.

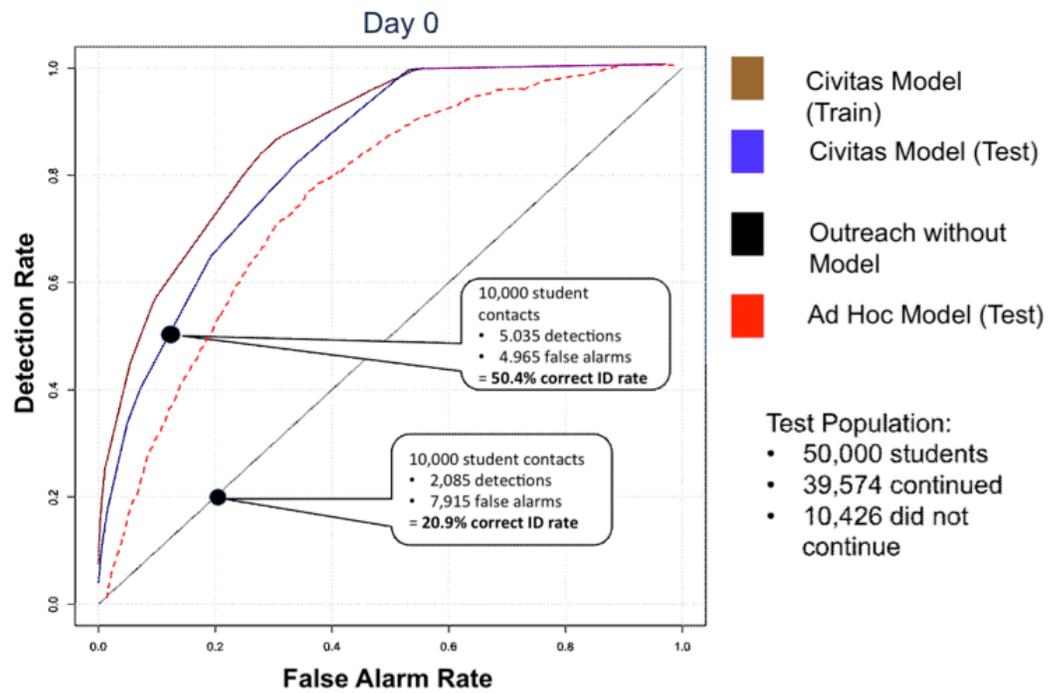


Figure 1.4. The day-0 ROC curves for the final train/test models using data-availability segmentation and clustering, an ad hoc model, and the random chance line.

Outcome Performance

How you bring data to the front lines of learning – e.g., to advisors and faculty – has a significant impact on the effectiveness of these efforts. Modality, timing, visualization, and operational tools matter.

January 2014. The test group outperformed the control group in persistence by 120 basis points. However, the p -value was 0.22, not reaching statistical significance. Institution C found these results to be promising and in the following term made plans to operationalize a daily prediction report.

March 2014. There was no measurable impact on persistence in the March 2014 term. Theories as to why there was no improvement focused on the operational complexity of managing a nightly report and distributing assignments to advisors in a timely fashion. Development of an application interface for advisors was underway and became the highest priority for the next pilot.

May 2014. Among new students, persistence rates from the May term into the July term were 762 basis points greater for test group than the control group. This result had a p -value of 0.02, reaching statistical significance at a confidence level of 98%. There was no measurable improvement for students past the ninth term of enrollment. Positive, statistically significant improvement was seen for students in their second until seventh term, into their eighth term.

In addition, intervention approaches were analyzed by student persistence probability and also by terms completed. For “Low” and “Moderate Persistence Probability” students, phone calls where the advisor “spoke to” the student were the most effective intervention approach. However, for “High Persistence Probability” students, “spoke to” was only slightly better than an email intervention.

In reviewing the intervention data by terms completed, for early term students, phone calls where the advisor spoke to the student were the most effective intervention. Conversely, for students with greater than ten terms completed at the institution, email appears to be the best initial intervention. However, if the student does not respond to the email, a phone call follow-up became the most effective approach.

Final Discussion and the Road Ahead

Each of these case studies involved institutions doing the work of developing deep insight analytics capacity and deploying action analytics strategies. From the results of these and other projects across our institutional cohorts, we point to the following observations as keys to leveraging these strategies to impact student learning and completion work:

Insight analytics that are developed using institution-specific data sources – particularly student-level SIS and LMS data – are vital to understanding student flow, as well as targeting and personalizing intervention and outreach. In short, there is not a global predictive model that works across institutions with any level of accuracy. You need to “turn the lights on” in your institution.

- The inclusion of additional data streams in insight analytics work can add value in better understanding student flow and targeting outreach.
- Adding ongoing activity data from students improves the performance of model predictive power.
- Bringing insight analytics together with action analytics is essential to “moving the needle” on student success. Better precision of the models helps target outreach and improve impact of instruction and advising support.
- Trying and testing action analytic outreach is a must. The work of iterating on outreach, what some in our community are calling intervention science, results in the best outcomes. There are no silver bullets, and tuning outreach to a unique student population is key. Put simply, the predictive models are just the beginning of the work.
- How you bring data to the front lines of learning – e.g., to advisors and faculty – has a significant impact on the effectiveness of these efforts. Modality, timing, visualization, and operational tools matter.

We summarize these findings in a simple framework we call the challenge of the four rights: (a) building the right infrastructure to (b) bring the right data to (c) the right people in (d) the right way. Importantly, the right way may be the most difficult aspect, because it includes how we visualize data, operationalize interventions and outreach, choose modalities, provide real-time feedback, and test the timing of interventions and outreach. In many ways, this is the art and science of analytics initiatives in higher education. Moreover, we need to ensure that we take security, privacy, and especially the impact of unintended consequences seriously. Indeed, data brought the wrong way to at-risk students – e.g., a flashing red indicator that in essence tells them that they are destined to fail – might do great damage to a population we care about a great deal (Stevens, 2014). That is why the trying and testing of outreach as a discipline is key here.

Going forward, the work of the Civitas Learning community will be focused on how we continue to bring together the best of insight and action analytics to help students learn well and finish strong on higher education pathways. Much is to be done, and much is to be learned. But as the field of analytics continues to take shape in higher education, there is clearly great promise. However, learning together will be essential.

Moreover, we need to ensure that we take security, privacy, and especially the impact of unintended consequences seriously. Indeed, data brought the wrong way to at-risk students – e.g., a flashing red indicator that in essence tells them that they are destined to fail – might do great damage to a population we care about a great deal.

References

- Blumenstyk, G. (2014, September 15). Companies promise “personalized education.” *The Chronicle of Higher Education*. Retrieved from <http://chronicle.com/article/Companies-Promise/148725/>
- Christakis, N., & Fowler, J. (2007). The spread of obesity in a large social network over 32 years. *The New England Journal of Medicine*, 357(4), 370–379.
- Civitas Learning Space. (2014). Retrieved from <http://www.civitaslearningspace.com/>
- Complete College America. (2014). Retrieved from <http://completecollege.org/>
- Crosta, P., & Kopko, E. (2014). Should community college students earn an associate degree before transferring to a four-year institution? *CCRC Working Paper No. 70*.
- Deeper Learning Initiative. (2014). *The William and Flora Hewlett Foundation*. Retrieved from <http://www.hewlett.org/programs/education/deeper-learning>
- Education Initiative. (2014). *The Kresge Foundation*. Retrieved from <http://kresge.org/programs/education>
- Fain, P. (2014, March 27). Predicting success. *Inside Higher Ed*. Retrieved from <https://www.insidehighered.com/news/2014/03/27/civitas-brings-completion-oriented-big-data-community-colleges-and-universities>.
- Fisher, R. A. (1954). *Statistical methods for research workers*. Edinburgh, UK: Oliver and Boyd.
- Higher education. (2014). *Obama administration 2020 goals*. Retrieved from <http://www.whitehouse.gov/issues/education/higher-education>
- Hill, J., & Su, Y.-S. (2013). Assessing lack of common support in causal inference using Bayesian nonparametrics: Implications for evaluating the effect of breastfeeding on children’s cognitive outcomes. *The Annals of Applied Statistics*, 7(3), 1386-1420.
- Kim, J. (2014, January 26). 8 questions for the co-founders of Civitas Learning. *Inside Higher Ed*. Retrieved from <https://www.insidehighered.com/blogs/technology-and-learning/8-questions-co-founders-civitas-learning>
- Lumina Foundation strategic plan 2013-2016. (2013). *Lumina Foundation*. Retrieved from http://www.luminafoundation.org/goal_2025.html
- Milliron, M. (2013, December 9). Confusing signals. *Inside Higher Ed*. Retrieved from <https://www.insidehighered.com/views/2013/12/09/essay-real-meaning-debate-over-early-notification-system-students>
- Milliron, M. D., & Rhodes, R. M. (2014). What do we expect? A regional education ecosystem that works. *The Texas Lyceum Journal*, 15-22.
- Phan, N., Xiao, X., Dou, D., Piniewski, B., & Kil, D. (2014). Analysis of physical activity propagation in a health social network. *ACM International Conference on Information and Knowledge Management*, Shanghai, China.
- Postsecondary success strategy overview. (2014). *Bill and Melinda Gates Foundation*. Retrieved from <http://www.gatesfoundation.org/What-We-Do/US-Program/Postsecondary-Success>
- Rees, J. (2014, September 3). September Pioneer Summit 2014. *Civitas Learning Space*. Retrieved from <http://www.civitaslearningspace.com/september-pioneer-summit-2014/>
- Stevens, M. L. (2014). An ethically ambitious higher education data science. *Research & Practice in Assessment*, 9(2), 96-98.
- Texas Student Success Council. (2014). *Educate Texas*. Retrieved from <http://www.edtx.org/postsecondary-access-and-success/postsecondary-success/texas-student-success-council/>
- Thornburgh, C., & Milliron, M. D. (2013, December). Driving analytics to the front lines of education. *CIO Review*. Retrieved from <http://www.cioreview.com/magazine/Driving-Analytics-to-the-Front-Lines-of-Education-FKCU497192190.html>

Appendix

Deep Dive on Some of the Data Science behind Insight and Action Analytics

Overview of Insight and Action Analytics

Extracting actionable insights from data requires a complementary fusion of (a) extraction of insightful derived features, (b) ranking and optimization of features in a hierarchical learning network to accommodate a diverse collection of data footprints of students, and (c) visual analytics to surface complex information in an intuitive way.

Feature extraction is a continuous quest to encapsulate and bring to light useful information that can be acted upon. In this Appendix, we show examples of various insights in one-, two-, and multi-dimensional plots in an increasing order of complexity. Figure 1 shows a few examples of insightful features in marginal class-conditional densities. The probability density functions (PDFs) in green and orange are $p(x/y=\text{persist})$ and $p(x/y=\text{not persist})$, respectively, where x = student feature and y = student success outcomes or classes in classification.

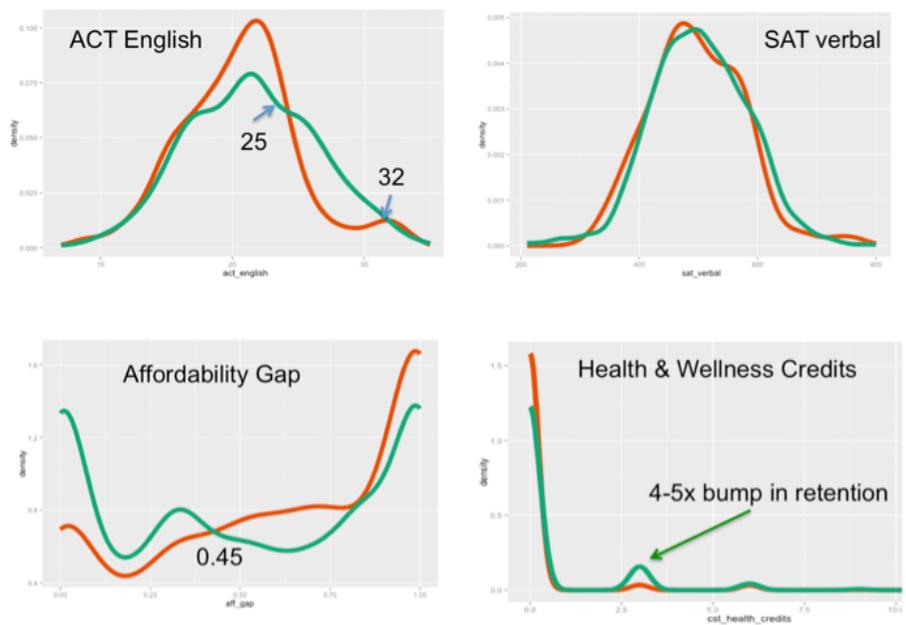


Figure 1. Examples of insightful features. With the exception of plot from the ISSM model, the rest are derived from persistence prediction models.

The ACT English plot is interesting in that SAT Verbal was not a strong predictor of persistence. When we probed deeper, we learned that this institution places a heavy emphasis on writing in all their courses. ACT English measures writing skills while SAT Verbal does not.

Another example is that the affordability gap (AG) shown in the lower left-hand corner is more insightful than raw financial aid amount since AG measures the ratio of financial aid to tuition owed. Such a plot can provide insights into how to allocate Pell Grant financial aid to improve persistence of Pell Grant recipients.

The Health & Wellness plot shows that students who take one health & wellness course as an elective persist at a much higher rate. While this observation does not imply causation, it can lead to an interesting research question and experiment design.

The class-conditional feature PDFs compare incoming student success rates as a function of the percentage of single-parent households in zip codes students come from. An actionable implication here is that if an incoming student has a high risk of not persisting and is from a high single-parent household area, she may be a prime candidate for a mentorship program, especially if a mentor has a similar background in the beginning, but has been academically successful with good social skills.

In certain situations, a combination of more than one feature brings out more meaningful insights. Pathway analysis has generated a lot of interest, especially for community college (CC) students (Crosta & Kopko, 2014). Figure 2 shows clearly that the probability of earning a bachelor's degree reaches a peak at around 60 credit hours. That is, CC students who earn AA/AS degrees improve their probability of earning bachelor's degree by more than 10% from the baseline trend for all transfer students.

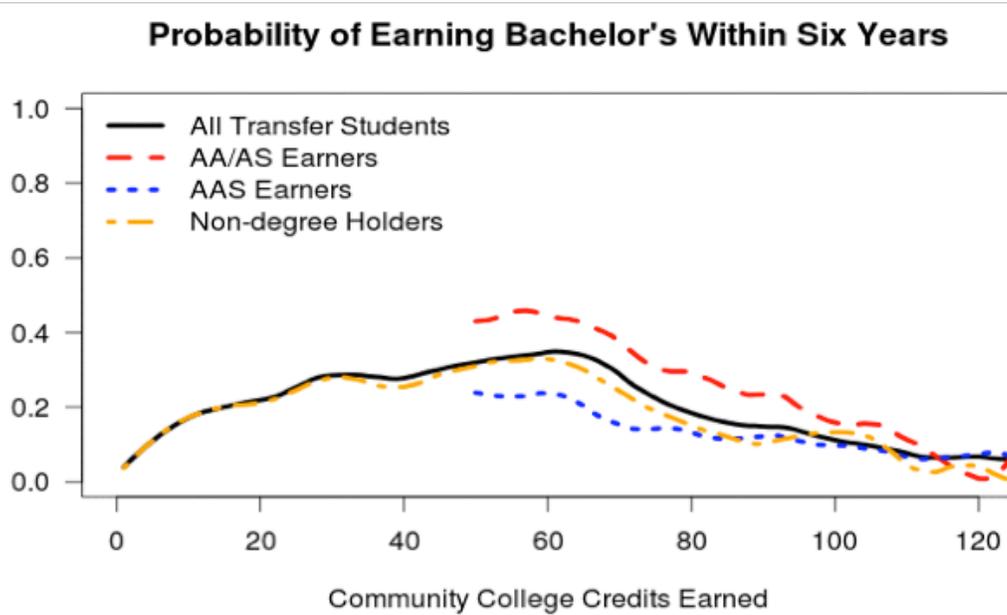


Figure 2. College pathway analysis (Crosta & Kopko, 2014).

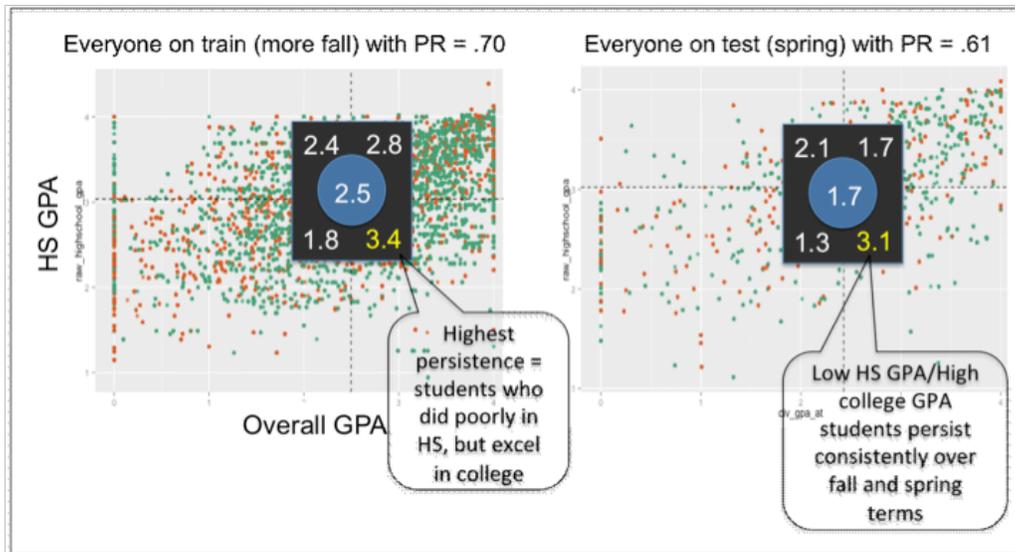


Figure 3. The 2 x 2 scatter plots over high school GPA and community college GPA paint an interesting picture. The five numbers in the centroid (50%-50% line) represent the ratio of the number of students who persist to that of students who do not for all and each of the four quadrants. Persistence rate drops significantly in spring, in part due to high-performing students transferring out.

We are currently federating data between 2- and 4-year schools, where the 2-year institutions serve as feeder schools to the 4-year institutions, so that we can do more thorough investigation into optimal transfer pathways and how to apply personalized interventions to students who are likely to benefit by finishing AA/AS degrees at community colleges.

In general, students with high CC GPA in the spring term tend to transfer out, which may suggest that advisors should target high-GPA students in the spring term to help them be better prepared by staying an extra year to earn AA/AS degrees. However, when we overlay another feature, high school GPA, a more interesting picture emerges as shown in Figure 3.

The 2 x 2 scatter plots use the same color code as in Figure 1. Each scatter point represents a student with color denoting the persistence flag (orange = not persist, green = persist). The number in the blue circle represents the ratio of those who persisted to those who did not. The four numbers along the edge depicts the same numbers in the four quadrants along the centroid.

The first observation is that the persistence rate (PR) is much lower in spring. The second key finding is that students with low high school GPA and high CC GPA (quadrant 4) tend to persist at a much higher rate than those with high GPAs in high school and CC, as well as their persistence rate being less dependent on seasonality. This finding alone can help advisors improve their targeting. Another example deals with the impact of scholarship on persistence as shown in Figure 4.

The left plot shows that merit scholarships given to students with high ACT scores are not as effective as those given to students with high high-school GPA. What is also interesting is that students who have high school GPA tend to persist at a higher rate than those with ACT scores. This shows the importance of multidimensional decision making by factoring into all key drivers of student success that depend on which segments and clusters they belong to in the hierarchical learning network based on data availability and clustering within each data-availability segment.

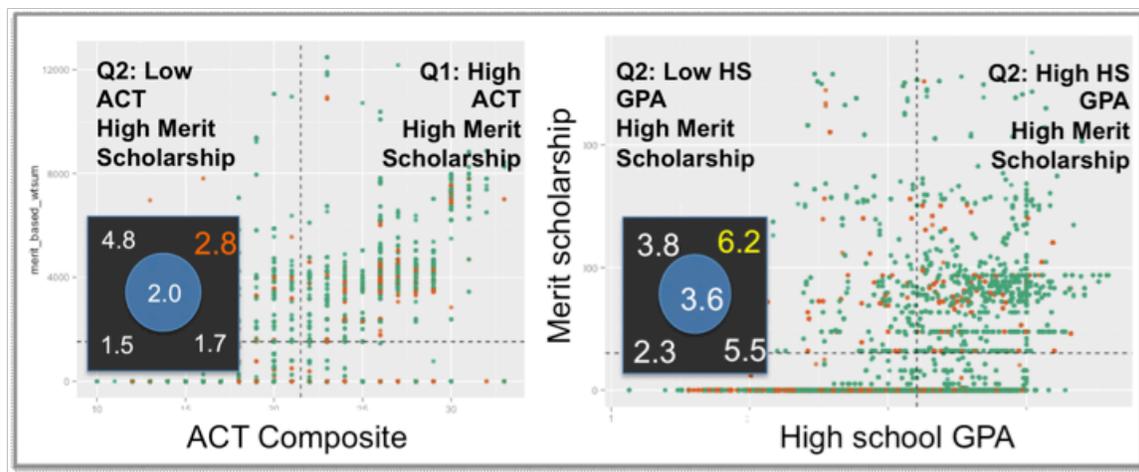


Figure 4. The impacts of scholarship on persistence.

Now we can extend the 2 x 2 concept indefinitely to provide insights with an arbitrary number of top features and/or at the segment/cluster level, where segments are determined based on available data footprints while clustering finds homogeneous groups within each segment, thus facilitating a hierarchical network view of the entire student population. Figure 5 shows the cluster heat map view. Columns and rows represent clusters and z scores (mean/standard deviation) of various attributes that characterize each cluster. The first two rows are population size (N) and persistence rate of each cluster. The rest of the rows represent various attributes, such as census household income, % of population with BA degree or higher based on census, age, cumulative GPA, the number of distinct 2-digit CIP codes in course work per term, etc. This quilt view extends much further in reality, while highlighting differences among the clusters based on color gradients across each row. Figure 5 shows 3 sets of clusters (low, medium, and high) grouped based on actual persistence rates. Table 1 compares and contrasts these performance-based clusters.

Furthermore, graph theories can be applied to understanding course pathways and the impacts of emerging influencers and cliques on helping other students succeed. Figure 6 shows a concurrent social graph and a time-varying series of student social networks over the course of a term.

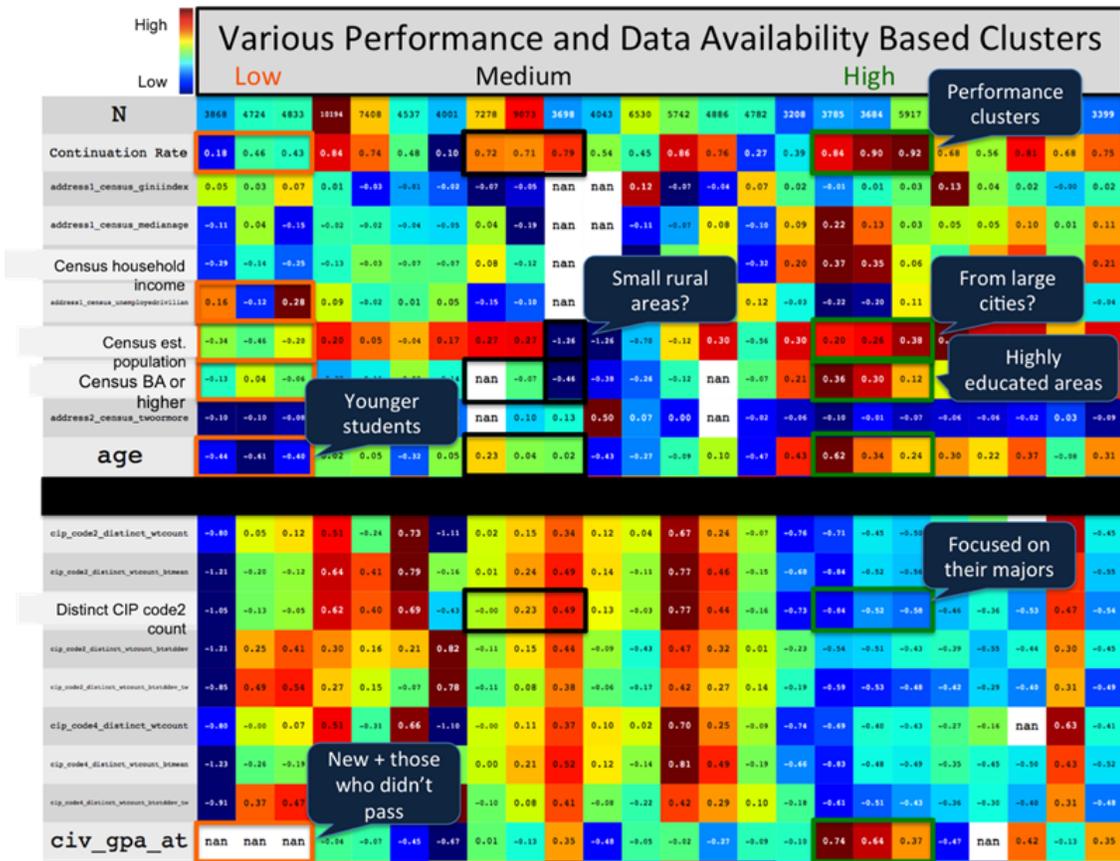


Figure 5. The cluster heap map view so that we can glean insights into how these clusters can be differentiated based on demographic variables, census-derived features, and top predictors. The white color indicates that the corresponding features and their associated raw data elements do not exist.

Table 1
A Comparison of the Three Performance-Based Clusters

Attribute	Low persistence	Medium persistence	High persistence
Persistence rate	~40%	~75%	~92%
Census household income	Low	Medium	High
Census estimated population	Small	Mixed	Large
% of residents with BA or higher degrees	Low – med	Low	High
Student age	Young	Middle	Mature
Distinct CIP code2 count	Low to med	High	Low
Cumulative GPA	N/A	Medium	High
Financial aid	Pell + some loan	Some Pell, little loan	High loan, little Pell
Terms completed	None	Most	Middle

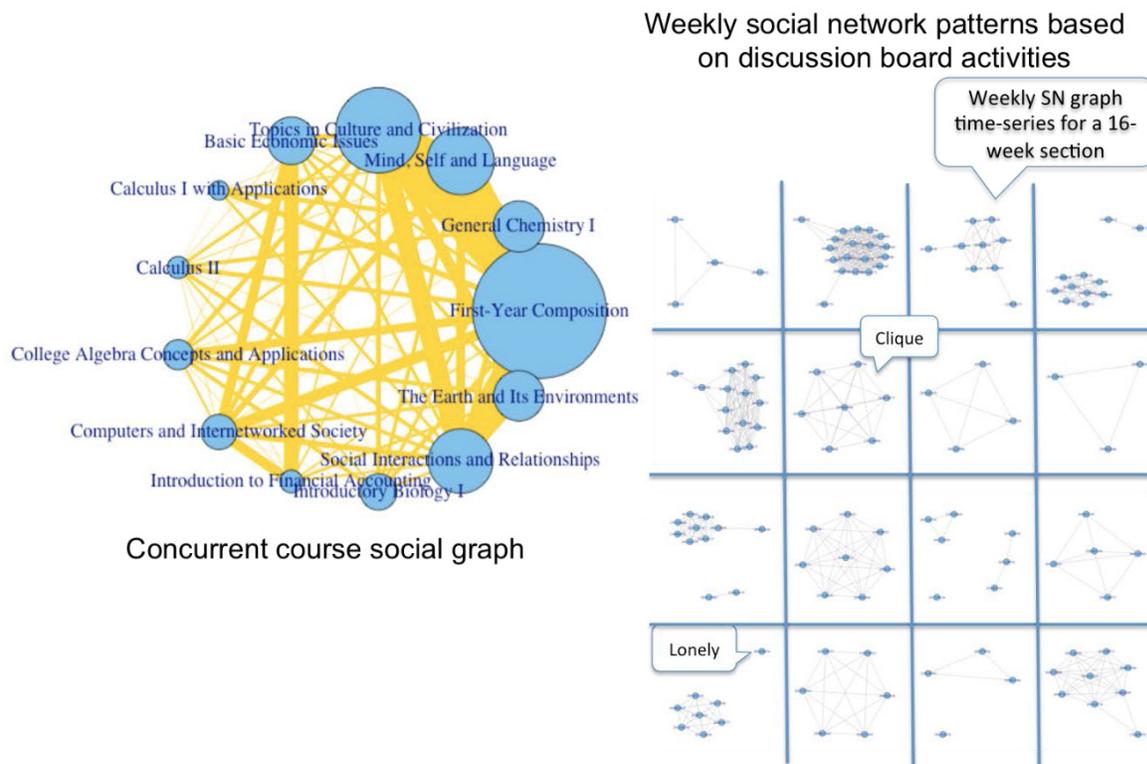


Figure 6. Course social graph and social network dynamics throughout a term.

The concurrent course social graph shows what courses are being taken together with the vertex size proportional to enrollment. The thickness of edges between courses is proportional to how frequently the connected courses are taken together. This allows us to investigate students' course-taking behaviors and toxic/synergistic course combinations by melding successful course completion predictions, propensity score matching by creating test and matching control groups, and explicitly incorporating students' course-taking patterns. The same analysis can be extended to course pathways over multiple terms to help us glean insights into the paths taken by successful vs. less successful students.

The same concept applies to social network analysis. Christakis and Fowler (2007) found that obesity spread through one's social network. Phan et al. (2014) applied the concept further by identifying emerging influencers and then studying their influence on connected pilot participants as a function of time to quantify how good health behaviors can be spread through peer-to-peer nudging, discussion board, and sharing of pedometer data through games. We plan to apply similar methodologies in student social networks so that we can work with faculty in facilitating students helping other students under faculty nudging. Our preliminary work indicates that a few social network features are statistically significant in predicting successful course completion and persistence.

Examples of Action Analytics

Action analytics can be most effective when actionable insights are brought to frontline people and their intervention details are captured in database tables for an integrated predictive and intervention science research. In principle, the predictive science provides insights into who is at risk, when the right moment for engagement or intervention is, and what intervention will be effective down to an individual student level. Intervention science works in concert with predictive science to provide foundational data for computing intervention utility, which in turn becomes the basis for intervention recommendation.

Intervention science data comes from encoding all facets of interventions – type, delivery modality, messaging attributes, business rules for intervention (who, when, and why), and primary/secondary endpoints for outcomes. Intervention science analytics encompass experiment design, power analysis, propensity score matching (PSM), Bayesian additive regression trees (Hill & Su, 2013), predictive modeling, and predictive ratio analysis. All these methods can shed scientifically rigorous insights into what interventions work or do not for which groups of students under what context. Figure 7 shows our overall framework for intervention science.

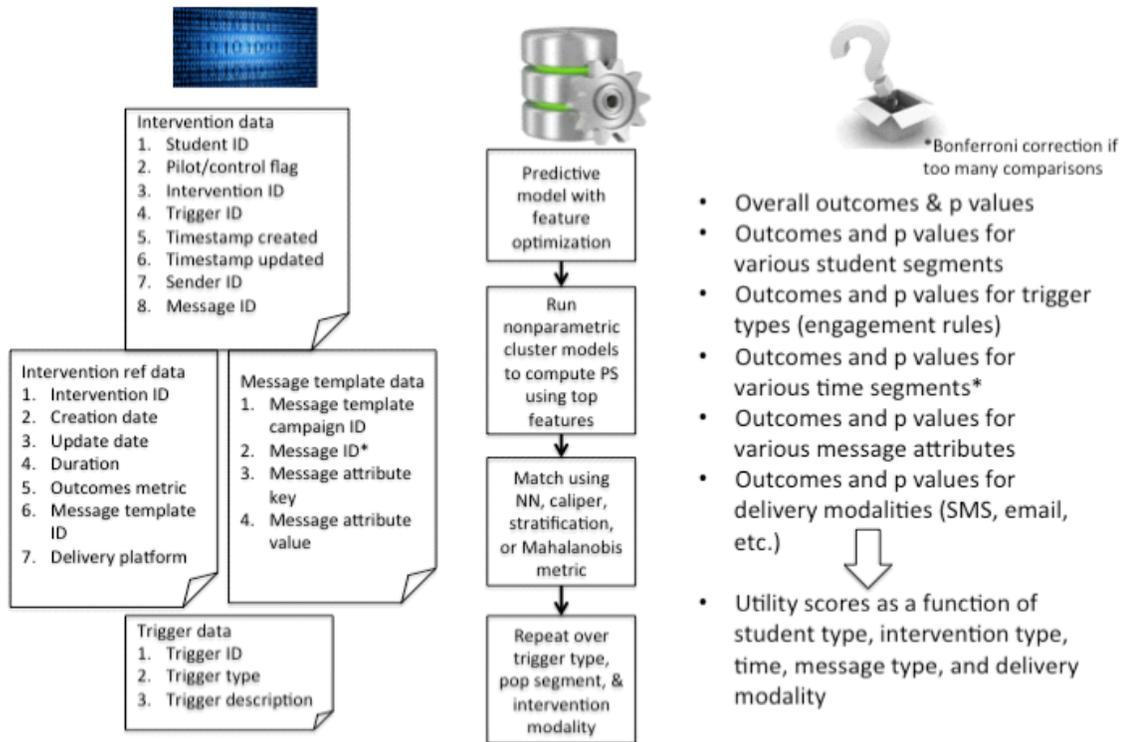


Figure 7. Our intervention science framework that leverages both predictive models and drill-down outcomes analytics to provide insights into intervention efficacy.

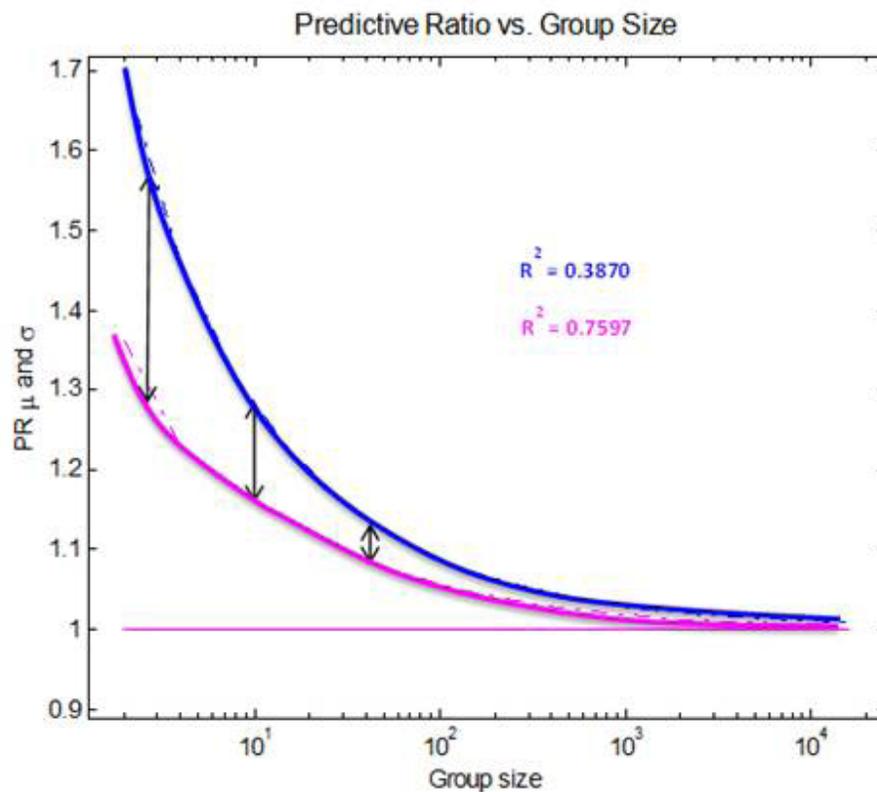


Figure 8. The more powerful the model is measured by R2, the smaller the standard deviation in predictive ration (PR) is at various group sized, leading to greater statistical power, i.e., a lower minimum detectable threshold in outcomes differences between pilot and control.

Action analytics apps surface to frontline prediction scores and key risk drivers at an individual student level. They also provide real-time feedback on intervention efficacy by showing how student engagement scores, prediction scores, and early enrollment statistics are changing for the pilot group in comparison to the control group. We select students in the control group through randomization and/or PSM prior to the commencement of a pilot.

In order to maximize statistical power in outcomes analysis, we apply hierarchical modeling techniques based on data availability, where a model is instantiated at the segment level. For each segment, we use the model's top features in PSM. The more predictive the models are using these top features, the greater the statistical power is. Figure 8 demonstrates that the higher-performance model in magenta exhibits a lower standard deviation curve for predictive ratio at all group sizes. Furthermore, we augment PSM with prediction-score matching such that matching cohorts have similar PDFs in propensity and prediction scores.

In summary, action analytics take risk predictions as an input in order to identify when to apply which interventions to which students. Once interventions are applied, we use various primary and secondary endpoints to investigate the efficacy of interventions as a function of engagement business rules, population segments, and intervention modalities. We provide real-time feedback for advisors and faculty by pointing out how their interventions are affecting feature and prediction score PDFs since human factors also play such an important role in affecting intervention outcomes. This information becomes the foundation of action analytics and intervention science.