WHITE PAPER

Predictive Analytics in Higher Education

Data-Driven Decision-Making for the Student Life Cycle

January 2013

EDUVENTURES

Eduventures, Inc. | 101 Federal Street, 12th Floor | Boston, MA 02110 | www.eduventures.com

Table of Contents

Introduction

Broadly defined, predictive analytics is "an area of statistical analysis that deals with extracting information using various technologies to uncover relationships and patterns within large volumes of data that can be used to predict behavior and events." This definition, which was provided by Educause, is relatively new and will serve as the framework for this white paper.

Higher education is a relatively late adopter of predictive analytics as a management tool. Predictive analytics has been used in other industries for many years, especially in the area of assessing consumer behavior. For instance, automobile manufacturers and dealers use predictive analytics to assess the likelihood of a customer who leases a car to either purchase the vehicle or choose to lease a new vehicle at the end of the lease. Using predictive analytics, BMW might extend an offer for a no-penalty early lease termination to select customers who the data suggests are likely to move to Mercedes Benz or Audi, if they agree to a new BMW lease. In this scenario, offers would not be made to customers who the data suggest are predisposed to continue with BMW; instead, they are designed to generate repeat business from those most likely to defect.

In a similar way, colleges and universities can deploy predictive analytics to determine which students are most at risk for attrition and – armed with deep, historical data – craft segment-specific retention strategies designed to compel them to persist toward degree completion.

Eduventures believes that predictive analytics, while by no means a "silver bullet" to solve all of higher education's problems, is an important tool that can positively impact efficiency and effectiveness across the academy. To better understand predictive analytics, Eduventures undertook research that includes interviews with practitioners from a variety of institutions, including public and private four-year universities and a community college. We sought a selection of interviewees who are analytics thought-leaders or who at least had experience in its implementation.

This white paper is designed to explain predictive analytics, followed by a look at how it can impact activity at the highest levels of institutional management. We provide examples of how predictive analytics has been used at a variety of institutions, including a review of its potential pitfalls and benefits. Citing interviews with practitioners, this white paper provides concrete examples of how predictive analytics has led to measurable performance improvements. Finally, we close with a call to action to all colleges and universities to consider building predictive analytics into their toolbox of techniques that inform and enable evidence-based decision-making.

Predictive Analytics Defined

When applied to higher education, predictive analytics can help institutions accurately predict student behaviors – notably in the areas of learning outcomes, recruitment, and retention. For instance, by analyzing historical data, predictive analytics can inform an institution as to which applicants are most likely to enroll and, later in the student life cycle, which are likely to persist and graduate. Armed with these data, institutions can intervene with those who show signs of trouble, in real time, before it is too late to effectively intervene.

Broadly defined, predictive analytics is "an area of statistical analysis that deals with extracting information using various technologies to uncover relationships and patterns within large volumes of data that can be used to predict behavior and events." (Source: Educause)

For the purposes of this white paper, we look at predictive analytics as a valuable tool with which to engineer positive change throughout the student life cycle. As the cost to recruit a student rises, it becomes ever more important to retain students until they graduate, which will:

- Improve student learning outcomes.
- Improve retention and graduation rates.
- Improve the institutional return on investment (ROI) on recruitment costs.
- Increase operational efficiency.
- Help the institution demonstrate success in a key area of focus for accrediting agencies and the Federal government.
- Demonstrate positive efforts to other important entities (e.g., state legislatures that allocate funding to public colleges and universities).

Beyond the student life cycle, predictive analytics can be used across the academic enterprise – from advancement ("What's the likelihood of an alumni subset making planned gifts or attending homecoming?"), to residential life ("If we make $x investment in dorm upgrades, will we recoup this through longer stays and higher rates?"), and to academic affairs ("Assuming current recruitment and retention rates, how many adjunct faculty members will we need in the College of Fine Arts in 2015?").

That said, for the purposes of this white paper, Eduventures will focus on the uses and considerable benefits of using predictive analytics to effect positive change in the student life cycle.

Why Predictive Analytics? Why Now?

A desire for stronger operations, as well as access to ongoing funding support and financial aid programs, is leading colleges and universities to change how they operate. With tuition far outpacing inflation, a $1 trillion cumulative student debt load burden that exceeds America's collective level of credit card debt, declining state funding, and six-year graduation rates that remain stubbornly below 60%, universities face mounting pressure from a variety of entities to find and operationalize more efficient business practices.

Higher education has been judged wanting in the court of public opinion as a result of these long-term challenges. While often called "the year of the MOOC,"[1] 2012 may well have been called "the year of growing discontent." From the cover of *Time* to numerous articles in *The New York Times*, the popular press has chronicled much that is perceived to be wrong with higher education, including ever-increasing cost, lack of accountability, and – perhaps for the first time – probing questions about whether a college education is worth the cost.

To explore how predictive analytics is being deployed in higher education, Eduventures conducted telephone interviews with six practitioners from across the spectrum of higher education, encompassing relevant experiences at the following U.S. institutions:

- American Public University System
- Arizona State University
- Baruch College
- Slippery Rock University
- Sinclair Community College
- Wichita State University

Our interviewees spoke often about how predictive analytics allows them to make informed decisions and, equally importantly, to make them faster. The ability to take prompt action can make all the difference between, for example, meeting the fall start goal and missing it by 10 students, particularly in an environment where 10 students can have a large impact on the finances of an institution.

[1] Massive open online course

Consider this scenario: A private non-profit institution misses its fall start by 10 students. At average student tuition of $28,500 per year, 10 students reflect a $285,000 deficit in that year. Extrapolated over the six years it takes many students to graduate, this seemingly minor miss of 10 students equates to $1,710,000 in lost revenue. In addition, beyond the enrollment management/recruiting effort, there are likely other implications – for example, fewer students may obviate the need for planned adjunct faculty hires. Similarly, residential life will have fewer "heads in beds" and, facing a significant loss of revenue, would need to increase efforts to encourage upperclassmen to consider another year in the dorms. Had this institution made the initial investment to build a predictive analytics function, it could have made early adjustments to avoid the recruiting shortfall.

Adding predictive analytics to the institutional management toolbox allows for a continuous learning loop in which analysis informs decisions. These decisions lead to outcomes that are then assessed and combined with updated data to make better-informed decisions. Getting it right the first time is unlikely, but making steady progress through iterations of data-informed decisions ultimately pays significant benefits.

Although gathering, analyzing, and interpreting data can be challenging, the undeniable payoff is the ability to make informed decisions that drive success, rather than leaving success to chance

Following is a review of our findings, including the myriad challenges with deploying predictive analytics, as well as some examples of the benefits of persevering to the point where predictive analytics generates actionable information.

Starting Point: Consistently Collecting Data

According to our interviews, an early and important challenge with predictive analytics actually precedes the use of analytics: gathering accurate data that can be dissected, analyzed, cross-referenced, and transformed to inform strategy. As colleges and universities have slowly adapted to technology over time, many have five, ten, or even more discrete systems collecting data. For instance, while the registrar may use a legacy mainframe system, enrollment management might use a combination of Banner and ApplyYourself, and academic departments may use the Blackboard learning management system. As one interviewee noted, data in so many sources leads to "multiple versions of the truth." It takes time, energy, and resources to either build bridges to get these disparate systems to communicate with one another or, alternatively, make a significant investment in a campuswide enterprise system.

Even if an institution's efforts to coalesce data are in their infancy, there is no reason that good, reliable data can't be collected right away – even if done in departmental silos. For this reason, we posit that the first step toward the myriad benefits of predictive analytics is to ensure that data are collected now. This is paramount because predictive analytics is based in large part on assessing historical patterns. Therefore, collecting data today will pave the road toward operational insights tomorrow.

Even institutions that haven't resolved this common data dilemma can begin setting the stage for predictive analytics in the future by systematically collecting data now. The underlying information upon which predictive analytical work will eventually be based is likely available today.

Campus administrators who can assure the reliability of their information will be well prepared to deploy more complex analytic efforts when all systems are in place as long as there is faith in the accuracy of the data. There are numerous off-the-shelf products and available consulting services that can be used to help gather, scrub, and share data and others that can be deployed to help with college-specific analytics.

Eduventures' scan of the marketplace suggests that one of the most well-known products is IBM's SPSS predictive analytics, although others are available. While every system has its merits and drawbacks, the dominant competitor in higher education predictive analytics has been no system at all. However, in an age of needing to do more with less, doing nothing is a less and less tenable option for most colleges and universities.

Predictive Analytics in Student Recruitment

Colleges and universities spend millions of dollars to recruit students. Contrary to expectations, the advent of the Internet did not obviate the need for costly printed viewbooks, postcards, and other marketing materials. In addition to these marketing collateral and the postage needed to send them, institutions spend large sums for admissions counselors to travel to college fairs, visit high schools, and host special events on campus and in other cities. They also spend money on sophisticated Internet marketing (e.g., paid search and search engine optimization) to capture attention from prospective students. This does not include the money spent on "tuition discounting," by which admissions leaders selectively allocate institutional aid to attract more students.

Although it can be difficult to quantify these costs, a recent report from Noel-Levitz suggests that the median per-student recruitment cost is $2,185. One interviewee for this study suggested that at her public university, the per-student acquisition cost is somewhat comparable, but also that competitive for-profit universities spend $5,300 to recruit each student. Another interviewee explained how his institution's sophisticated use of predictive analytics allows the admissions team to scientifically assess which inquiries are most likely to become applicants based on a variety of factors, including geographic location, anticipated college major, ethnicity, and source of first contact. Armed with this information, the admissions team scores each inquiry's likelihood of applying and gears its efforts and spending accordingly.

One interviewee explained how his institution's strategic plan called for enhancing the academic profile of its student body and increasing diversity, which are two conflicting goals. Using IBM SPSS predictive analytics, he studied large data sets on current students to determine that, of the 350 high schools from which the institution actively recruited, 31 typically yielded a higher proportion of the students he most wanted to enroll. Armed with these data, the admissions team redoubled its efforts in these 31 schools by hosting more special events, tripling the number of visits, and making special efforts to build relationships with guidance counselors. The outcome of this exercise in predictive analytics is impressive: His institution saw an 8% increase in applications from these 31 high schools and saw a 20-point increase in the applicants' SAT scores.

The same administrator used predictive analytics at the dawn of the most recent economic recession. By studying not only internal institutional data but also data from New York City public schools and the Consumer Price Index, he was able to model the expected impact of the recession on next fall's incoming class. With this advance information, he increased the size of the spring class to offset projected fall declines and saved $780,000. This compelling example raises an important point to consider relative to the use of predictive analytics: The institution is only one among many sources of relevant data. Examples of other information that can be used in predictive analytics include data from the U.S. Department of Labor, the U.S. Census Bureau, and even career-oriented sites such as Salary.com, instructive when predicting which industries will offer high-paying jobs.

Predictive Analytics in Student Retention

Despite decades of focusing on recruitment, colleges and universities have too often failed to retain the students they have enrolled previously. Improving retention performance even slightly can have significant, positive implications for a university's fiscal position.

Consider that the national average freshman-to-sophomore retention rate is approximately 75%. This means that about one-quarter of the students who started in fall 2012 will not return to their freshman institution in fall 2013; therefore, an institution that enrolled 5,000 freshmen for fall 2012 can expect to lose 1,250 of them. At a median acquisition cost of $2,185, this implies a loss of $2,731,250 from increased costs of acquisition alone. When factoring in five subsequent years of foregone tuition and fees for each of those 1,250 non-returning freshmen, the costs of poor retention are high. Clearly, any tool that will help colleges and universities make data-driven decisions about which students are likely to persist will, over time, lead to dramatic improvements in efficiency and effectiveness – not to mention helping students graduate!

In addition, the U.S. Department of Education is studying tying retention and graduation rates to access to Title IV financial aid. Education Secretary Arne Duncan plans to "reward and incentivize universities to build cultures around completion."[2]

With six-year graduation rates at 58%, on average, colleges and universities can ill afford – financially, politically, or in the court of public opinion – to continue to "lose" 42% of the students they enroll.

While multiple factors play a role in successful student recruitment, most agree that retention is more complex. For traditional-aged students, there are dozens of factors impacting their likelihood to persist, many of which cannot be measured (e.g., homesickness, missing a high school boyfriend or girlfriend, or simply being emotionally unprepared for the freedom of living away from home).

[2] Tanya Caldwell and Jacques Steinberg, "Education Secretary Targets Colleges With Low Graduation Rates," *The Choice*, June 5, 2012 (http://thechoice.blogs.nytimes.com/2012/06/05/education-secretary-targets-colleges-with-low-graduation-rates/)

The literature is rich with examples of predictable factors that play a role in student retention and which, when capably addressed, can influence persistence. In fact, many of these retention-related factors can be studied during the recruitment phase, raising an interesting question: Can the admissions team make decisions about which students to pursue based not only on their admissibility and likelihood to enroll but also on their likelihood to persist? The answer, according to many, is "yes," which raises a number of compelling opportunities.

Factors influencing both admissibility and likelihood of persisting range from the purely academic (e.g., high school GPA, SAT/ACT scores) to geographic (e.g., distance between home and campus) to financial (e.g., expected family contribution) to social (e.g., Is he an athlete?). By studying historical data resident in the university's database (or databases, as our research suggests is most often the case), institutions can build profiles of students who are most at risk of not persisting and develop steps to intervene in a timely manner. Alternately, these analytics may inform decisions about which students to not admit based on an elevated likelihood of dropping out.

Even more compelling is the volume and depth of data accessible once a student enrolls in the college or university. For instance, one institution correlates numerous risk factors to build an at-risk profile. By analyzing large volumes of data, this institution has determined that the following risk factors are among the most prevalent:

- Number of logins on the learning management system
- Level of self-confidence (assessed with diagnostic tools)
- Level of social integration into campus life
- Study skills (e.g., time management, prioritization)
- Declaration of a major

By aggressively monitoring these factors, the institution knows who is most at risk – and with whom to intervene. Said intervention steps often include so-called "intrusive advising," in which at-risk students must engage with an advisor in order to even register for the next term. The academic advising enterprise is integrally important to retention and is another area that calls for examination.

One institution we interviewed struggled for years to get to the point where its data are reliable. Now that this has been accomplished, the institution plans to use predictive analytics to:

- Inform student recruitment efforts as its home state sees a big demographic decline in traditional-aged students.
- Determine which academic majors hold the highest promise for career-focused students, and adjust its course and major offerings accordingly.
- Determine the long-term viability of certain majors.
- Make faculty hiring decisions based on predicted enrollments in highly desired majors.
- Link financial aid awards to increased likelihood of persistence, and make use of limited institutional aid to enroll those students identified as most likely to persist.
- Explore the benefits and drawbacks of increasing admissions standards (e.g., If the institution becomes more selective, how will this impact retention?).
- Inform future capital improvements (e.g., If future recruitment efforts focus on healthcare-related majors, should the institution build a new science building instead of renovating the student union?).
- Improve academic advising through the use of dashboards, early-warning systems, and enhanced communication between advisors and students.

On the other hand, simply having good data does not suffice; rather, the institution must be able to collect, analyze, and use the data to make informed decisions. Another critical contributor to success with predictive analytics in the academy is having appropriate staffing – in the Institutional Research Department or elsewhere – to manage and use the data.

Next Steps

Predictive analytics holds great promise for helping colleges and universities make evidence-based decisions about a wide variety of issues impacting not only the student life cycle (recruitment and retention) but also other areas as well. Our interviews suggest that the most important step a college can take is to institutionalize the notion of predictive analytics from the highest levels of administration, including the President and Provost. Without this top-level commitment, it is unlikely that institutions will have the will or the resources to build a campus-wide, fully deployed analytics effort.

Once this commitment is made and publicized, the next step is ensuring that the institution has the capacity to collect and disseminate reliable data. This is no small undertaking, but fortunately, there are many companies eager to offer solutions. One of the most appealing aspects of predictive analytics is that returns on the investments are quantifiable and clear.

Predictive analytics in higher education is still relatively new, and the barriers to successful adoption and deployment can be high. However, there are many compelling reasons to make the investment, including improving recruitment and retention performance, improving the lives of students, maximizing operational efficiencies, and answering calls for institutional and programmatic accountability.

This white paper was underwritten by the following sponsor:

About IBM Business Analytics

IBM Business Analytics software delivers data-driven insights that help organizations work smarter and outperform their peers. This comprehensive portfolio includes solutions for business intelligence, predictive analytics and decision management, performance management, and risk management.

Business Analytics solutions enable companies to identify and visualize trends and patterns in areas, such as customer analytics, that can have a profound effect on business performance. They can compare scenarios, anticipate potential threats and opportunities, better plan, budget and forecast resources, balance risks against expected returns and work to meet regulatory requirements. By making analytics widely available, organizations can align tactical and strategic decision-making to achieve business goals. For further information please visit ibm.com/business-analytics.

To request a call or to ask a question, go to: ibm.com/business-analytics/contactus. An IBM representative will respond to your inquiry within two business days.

About Eduventures

Eduventures has worked in the education sector for 20 years. Today, Eduventures partners with more than 300 institutions of higher education to deliver the research and strategic guidance necessary to support effective decision-making on campus. By supporting each of our clients in the broad challenges they face on campus, we have become the leader in higher education research and consulting.

Please contact us at contact@eduventures.com or (857)221-9815 if you would like more information about how we can help you.

EDUVENTURES